

124425

STIC-Biotech/ChemLib

From: Mertz, Prema
Sent: Sunday, June 13, 2004 11:56 AM
To: STIC-Biotech/ChemLib
Subject: 10/009,557

Please search SEQ ID NO:34 with DNA databases.

Please search SEQ ID NO:9 with protein databases.

Thanks.

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
4D81 Remsen Bldg Mailbox 4C70
US Patent & Trademark Office
(571) 272-0876

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 6/16
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:58:40 ; Search time 45 Seconds
(without alignments)
2369.892 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYVLSPEVFIIQLLFIQAI.....PIISYFCSLGCYVNSSDMLK 338

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodeo.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvivirus.*
- 16: sp_bacteriap.*
- 17: sp_arctean.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284.5	16.3	428	5 Q9VKU0	Q9VKU0 drosophila
2	283.5	16.2	428	5 Q24386	Q24386 drosophila
3	247	14.2	366	11 Q8BX06	Q8BX06 mus musculu
4	247	14.2	403	4 Q96CX6	Q96CX6 homo sapien
5	244	14.0	311	16 Q8F116	Q8F116 leptospira
6	233.5	13.4	498	16 Q8F3F9	Q8F3F9 leptospira
7	227	13.0	423	16 Q8F7S1	Q8F7S1 leptospira
8	225.5	12.9	507	5 Q9N3F2	Q9N3F2 caenorhabdi
9	225	12.9	633	16 Q8F3F6	Q8F3F6 leptospira
10	224.5	12.9	452	16 Q8F118	Q8F118 leptospira
11	222.5	12.8	428	16 Q8F3F8	Q8F3F8 leptospira
12	218	12.5	679	5 Q61967	Q61967 caenorhabdi
13	216	12.4	426	16 Q8F213	Q8F213 leptospira
14	215	12.3	289	13 Q7ZSY7	Q7ZSY7 xenopus lae
15	212.5	12.2	1537	4 Q96NW7	Q96NW7 homo sapien
16	211.5	12.1	377	16 Q8F119	Q8F119 leptospira

ALIGNMENTS

RESULT 1

ID	Q9VKU0	PRELIMINARY;	PRT;	428 AA.
AC	Q9VKU0;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	IRR47 protein (GH06740P).			
GN	IRR47 OR C6098.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.N.,			
RA	Jodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris X.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwar C.,			
RA	Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,			

Q8imr8 drosophila
Q9vbe4 drosophila
Q8imr9 drosophila
Q8qcs9 drosophila
Q8tni4 methanosarc
Q9425 caenorhabdi
Q8cte7 mus musculu
Q9v900 drosophila
Q9v780 drosophila
Q8co29 mus musculu
Q8f117 leptospira
Q9nir8 dictyostell
Q8f212 leptospira
Q8f2b3 leptospira
Q9d5q5 mus musculu
Q9d2f4 mus musculu
Q9p977 candida alb
Q9f411 candida alb
Q9v4c4 homo sapien
Q86qt0 drosophila
Q7z5f7 homo sapien
Q7z5f6 homo sapien
Q86q87 drosophila
Q9nir8 candida alb
Q9esw8 homo sapien
Q9nw37 homo sapien
Q9w2u2 drosophila
Q8f3f4 leptospira

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003629; AAF52970.1; -
DR EMBL; AY058306; AAL13535.1; -
DR FlyBase; FBgn0010398; Lrr47.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 4.
DR SMART; SM00369; LRR_TYP; 1.
SQ SEQUENCE 428 AA; 47961 MW; AED96C86CA5B28D CRC64;

Query Match 16.3%; Score 284.5; DB 5; Length 428;
Best Local Similarity 28.7%; Pred. No. 3.8e-16;
Matches 96; Conservative 53; Mismatches 135; Indels 51; Gaps 10;

QY 15 LFTQAISSSLKGFSLAMRLAHRG-----CNVDTPVSTLTPTVKTSSEPFNKTKMVTSKD 69
Db 91 LLIKCDPIQLKGFLOTILKMGDKDALKLNRLNNAATAIPQKACP-----QVRNVIKSE 146
QY 70 YPLSKNFPYSLEHLOTSYGLVDMVMCLSKLRKLDLSHNHIKLPATIGLIHLQEL 129
Db 147 YPI-KGFPRTLSKLTNNSQLVKLSFEICTRLNLTCLDVSNGKLKIPSELGRL-PLTSL 204
QY 130 NLNDNHLESFS-----VALCHSTLQKLSRLSDLSKNKIKALPVQFCQLOELNKLDD 182
Db 205 HLGNNLLGTQNDWCWLRGTKLC-----QSLGELDLSGNGLTYPPLVVKFESLVSLNN 259
QY 183 NELIQPCKIGQLINLRFLSAARNKLPFLPSEFNRNLSLEYDLFGNTFE-----QPKV 235
Db 260 NLLSLRPFPAIRKALRKLVCNSELSPSAVEDLRIDLDVWGNCFKFNDAQAQMY 319
QY 236 LPVVKLOAPLTLESSARTILNRPYGSIIIPFHLCCQDLDTAKICVCGRFCLN 288
Db 320 LQKAASNPQWLGLGARAVDKYMLPLSAGSIIPAVLIDIREAPRCPCGELCYAQRKEDL 379
QY 289 -----NSFIQGTITNNL-----HVAHTVVLVDN 312
Db 380 FORVQPKFISVKNLTYSREHQIYADVVLCD 411

RESULT 2
Q24386
ID Q24386 PRELIMINARY; PRT; 428 AA.
AC Q24386;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LRR47 precursor.
GN LRR47 OR CG5098.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94289476; PubMed=8018718;
RA Ntwasa M., Buchanan S.G.S.C., Gay N.J.;
RT "Sequence and expression of LRR47, a novel embryonic leucine rich
RT repeat protein of *Drosophila*;"
RL Biochim. Biophys. Acta 1218:181-186(1994).
DR EMBL; X75760; CAA53387.1; -
DR PIR; S45361;
DR FlyBase; FBgn0010398; Lrr47.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 4.
DR SMART; SM00369; LRR_TYP; 1.
KM Signal.
FT SIGNAL.
SQ SEQUENCE 428 AA; 47856 MW; CECEL29ALD7F8F7F CRC64;

Query Match 16.2%; Score 283.5; DB 5; Length 428;
Best Local Similarity 27.7%; Pred. No. 4.7e-16;
Matches 92; Conservative 55; Mismatches 140; Indels 45; Gaps 8;

QY 15 LFTQAISSSLKGFSLAMRLAHRG-----CNVDTPVSTLTPTVKTSSEPFNKTKMVTSKD 69
Db 91 LLIKCDPIQLKGFLOTILKMGDKDALKLNRLNNAATAIPQKACP-----QVRNVIKSE 146
QY 70 YPLSKNFPYSLEHLOTSYGLVDMVMCLSKLRKLDLSHNHIKLPATIGLIHLQEL 129
Db 147 YPI-KGFPRTLSKLTNNSQLVKLSFEICTRLNLTCLDVSNGKLKIPSELGRL-PLTSL 204
QY 130 NLNDNHLESFS-----VALCHSTLQKLSRLSDLSKNKIKALPVQFCQLOELNKLDD 182
Db 205 HLGNNLLGTQNDWCWLRGTKLC-----QSLGELDLSGNGLTYPPLVVKFESLVSLNN 259
QY 183 NELIQPCKIGQLINLRFLSAARNKLPFLPSEFNRNLSLEYDLFGNTFE-----QPKV 235
Db 260 NLLSLRPFPAIRKALRKLVCNSELSPSAVEDLRIDLDVWGNCFKFNDAQAQMY 319
QY 236 LPVVKLOAPLTLESSARTILNRPYGSIIIPFHLCCQDLDTAKICVCGRFCLN 290
Db 320 LQKAASNPQWLGLGARAVDKYMLPLSAGSIIPAVLIDIREAPRCPCGELCYAQRKEDL 379
QY 291 -----FIQGTITNNLHVAHTVVLVDN 312
Db 380 FORVQPKFISVKNLTYSREHQIYADVVLCD 411

RESULT 3
Q8BX06
ID Q8BX06 PRELIMINARY; PRT; 366 AA.
AC Q8BX06;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical leucine-rich repeat.
GN C330018J07RIK.
OS *Mus musculus* (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";


```
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR000437; Prok_lipoprot_s.
DR Pfam; PF00560; LRR; 14.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPRCTIN; 1.
KW Complete proteome.
SQ SEQUENCE 498 AA; 57222 MW; E6D43494FD13FA73 CRC64;

Query Match 13.4%; Score 233.5; DB 16; Length 498;
Best Local Similarity 41.1%; Pred. No. 1.1e-11;
Matches 65; Conservative 29; Mismatches 59; Indels 5; Gaps 4;

QY 100 LKSLRKLDLSHNHIKKLPATIGDLIHQLNLDNDNHLNLSFVSVALCHSTLQKLSRLDLSK 159
DB 70 LKSLRKLDLSHNHIKKLPATIGDLIHQLNLDNDNHLNLSFVSVALCHSTLQKLSRLDLSK 126
QY 160 NKIKALPVQFCOLQELKNLKLDDNELIOPCKIGQLINIRFLSAARNKLPFLPSEFRNL- 218
DB 127 NRIILPVEIGRLQNLQDGLYKKNKLTTPPKIGQLQNLQKLSLNSRLTALPKRIGQLK 186
QY 219 SLEYLDLFGNTE-QPKVLPVLIKQAPLTLLESSARTI 255
DB 187 NLIQTLDQNOFTIILPKRIGQLQNLQKLSLNSRLTALPKRIGQLK 224

RESULT 7
Q8F7S1 PRELIMINARY; PRT; 423 AA.
AC Q8F7S1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leucine-rich repeat containing protein.
GN LA0873.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011272; AAN48072.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 13.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.
SQ SEQUENCE 423 AA; 48695 MW; 91CB3DADB63C7DF0 CRC64;

Query Match 13.0%; Score 227; DB 16; Length 423;
Best Local Similarity 31.6%; Pred. No. 3.3e-11;
Matches 65; Conservative 44; Mismatches 89; Indels 8; Gaps 5;

QY 39 NVDTPVSTLPVKTSEFENFKTKWITSKK-DYPLSKNFPYSLEHLOTYSYGLVRVDMRM 97
DB 139 N-DNNKLTLPKKEIGQLQNLQELSLNKLISLPTETIEQLSKLNLDLNNHNETTVSKEV 198
QY 98 LCKSLRKLDLSHNHIKKLPATIGDLIHQLNLDNDNHLNLSFVSVALCHSTLQKLSRLD 157
DB 199 MLETTLENILRSNKLKTIPTKEIRQLSKLVMLTGNQLTSLPKR- -EQLQ-NLKTNL 255
QY 158 SKNKIKALPVQFCOLQELKNLKLDDNELIOPCKIGQLINIRFLSAARNKLPFLPSEFRN 217
DB 256 GNRFOILSVEILDELKNLELNLYNQLEFPKVEGQLSKLYSLYHNOITTLPEVITQ 315
QY 218 L-SLEYLDLFGNTEQPKVLPVLIKQ 242
DB 316 LPDLQELHLSGNKI---TILPKRILQ 338
```

RESULT 8

```
Q9N3F2 PRELIMINARY; PRT; 507 AA.
AC Q9N3F2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y54E10A.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wollam A., Becker M., Graves T., Hawkins M.;
RT "The sequence of C. elegans cosmid Y54E10A."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024810; AAF60763.1; -.
DR WormPep; Y54E10A.6; CE24450.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 5.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
KW Hypothetical protein.
SQ SEQUENCE 507 AA; 56414 MW; 98452F1FEC50591E CRC64;

Query Match 12.9%; Score 225.5; DB 5; Length 507;
Best Local Similarity 27.6%; Pred. No. 5.6e-11;
Matches 89; Conservative 40; Mismatches 114; Indels 79; Gaps 12;

QY 53 SEFENFKTKWITSKKDYPLSKNFPYSLEHLOTS-----YCGLVVRVDMRML-----C 99
DB 30 TEANLISQKRVFESMSQLNLLSLTGCSLNLSLSSIRSCSNLHMLVLPKNDLKQLPDVDFC 89
QY 100 LKSLRKLDLSHNHIKKLPATIGDLIHQLNLDNDNHLNLSFVSVA-----LCHST 147
DB 90 LPKLEPDLNLSHNLDALPASISKCNLESILNLSLNNLSNLSFSPDISNLSLHIFDAHNT 149
QY 148 LQK-----SLRSLDLSKKNKALPVQFCOLQELKNLKLDDNELIOPPKIGQL 195
DB 150 ISKIPASLTSNLSAKLHTIILSHNSEIIVPDSLSNLKQLKELKIDENKLDKVPVIAHL 209
QY 196 INLRFLSAARNKLPFLPSEFRNLSLE-----YLDI-----FGNTEQPKV----- 235
DB 210 PKLVLDISKN--CFSESFPKLANDKRGKLVNLAQKIGKPIGNSEEQKAPESGSP 267
QY 236 -LPVLIKQAPLTLLESSARTILHN-----RIPYGSIIIPHLQDLDTAKICVGRFLNS 290
DB 268 DPEV--PDAPLTV-----RTGIENLTWRHRHPSVEIRPVLC-----CVINNVDLNG 312
QY 291 FIOGTTNLSHVAHTVLDVN 312
DB 313 GDSFKKFIATQTRHASALCN 334

RESULT 9
Q8F3F6
```

```

ID Q8F3F6 PRELIMINARY; PRT; 633 AA.
AC Q8F3F6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative outer membrane protein.
GN LA2450.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE011413; AAN49649.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF00560; LRR; 23.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 633 AA; 72082 MW; 8C15836852DE2DDB CRC64;

Query Match 12.9%; Score 225; DB 16; Length 633;
Best Local Similarity 29.3%; Pred. No. 8e-11;
Matches 73; Conservative 43; Mismatches 89; Indels 50; Gaps 7;

QY 10 ILQLLFLQALSSSLKGLFSLAMLAHRCNVDPVSTLTPVKTSFEN-----FKTKMV 63
DB 91 VIVLQKLESD-----LSENRL-----VMLPNEIGRLQLQELGLYKNKI 132
QY 64 ITSCKDPLSKNFPYSLSLHLQ-SYGLVRVDMRLCLSKLRKLDLSHNHKKLPATIGDL 123
DB 133 TFPKEIGQLQ-----NLQTLNQDNQATLTPVEIGQLQNLKLNKRLTLVLPKEIGQL 187
QY 124 IHLQELANDNHLESFVSVALCH-STLQ-----KSLRSLDLSKNKIK 163
DB 188 QNLQTLNQDNQATLTPVEIGQLQNLQTLGSENQTLTPFKEIGQLQNLKLNKRLTLVLPKEIGQL 247
QY 164 ALPVQFQLQELKMLDNDNLIQPPCKIGQLINRFLSARANKLPFLPSFRNL-SLEY 222
DB 248 ALPKEIGQLKYLENLEUSENQTLTPFKEIGQLKXKLDLGLGRNQTLTPFKEIGQLKMLQ 307
QY 223 LDLFQNTPE 231
DB 308 LDLCYNQPK 316

RESULT 10
Q8F118 PRELIMINARY; PRT; 452 AA.
AC Q8F118;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Leucine-rich repeat containing protein.
GN LA3321.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB011492; AAN50519.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR Pfam; PF00560; LRR; 15.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.

```

```

SQ SEQUENCE 452 AA; 52245 MW; 153B25C528FA21AA CRC64;

Query Match 12.9%; Score 224.5; DB 16; Length 452;
Best Local Similarity 30.1%; Pred. No. 5.9e-11;
Matches 74; Conservative 38; Mismatches 85; Indels 49; Gaps 8;

QY 46 TLTPVKTSFENFKT-----KMWITSKDYPLSKNFPYSLSLHLQTSYCGLVRRVDMRLCL 100
DB 61 TTLPKEIGQLQNLQDLISFNSLTLPKEIGQLRN-----LQELDLSEFNSLTLPKEVGQL 116
QY 101 KSLRKLDSLHNHKKLPATIGDLIHLQELANDNHLESFVSVALCHSTLCKSLR-----S 154
DB 117 ENLQRLDLHQNLATLPMIEIGQLKMLQELNSKL-----TLTPKEIRQLRNLOE 167
QY 155 LDLSNNKIKALPVQFCOLQ-----ELKNLK-----LDDNELIQFPCK 191
DB 168 LDLSNNKLTLPKEIGQLQNLKTLNLTQTLTPKEIGQLQNLKTLNLTQTLTPKE 227
QY 192 IGQLINRFLSARANKLPFLPSFRNL-SLEYDLFGNTFEO-PKVLPIVKLQAPITLLE 249
DB 228 IGELQNLLEILVLRNRTALPKEIGQLQNLQDLHLQNLQTLTPKEIGQLQNLQDLHLQ 287
QY 250 SSARTI 255
DB 288 NQLTTL 293

RESULT 11
Q8F3F8 PRELIMINARY; PRT; 428 AA.
ID Q8F3F8;
AC Q8F3F8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative outer membrane protein.
GN LA2448.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011413; AAN49647.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF00560; LRR; 15.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 428 AA; 48954 MW; 0F7DD095D1B8A1C7C CRC64;

Query Match 12.8%; Score 222.5; DB 16; Length 428;
Best Local Similarity 31.8%; Pred. No. 8.2e-11;
Matches 74; Conservative 41; Mismatches 95; Indels 23; Gaps 8;

QY 39 NVDTPVSTLTPVKTSFEN-----FKTKWITSKDYPLSKNFPYSLSLHLQTSYCG--- 89
DB 100 NLSSNQLTLPKEIGQLKLNQDLIDYNRLTI-----LPIEIG---KLQNLQTLSSNQ 151
QY 90 LVRVDMRLCLSKLRKLDLSHNHKKLPATIGDLIHLQELANDNHLESFVSVALCHSTLQ 149
DB 152 LTLTPRESGKLENQELNSNQLTLPQEIQLQNLQTLNLSNQLTTFKEI---EGL 208
QY 150 KSLRSLDLSKNKIKALPVQFQLQELKMLDNDNLIQPPCKIGQLINRFLSARANKLP 209
DB 209 KNLQTLNSDNQLTLPKEIGQLQNLQTLNLSNQLTLPKEIGQLQNLQTLNLSNQLT 268
QY 210 FLPSFRNL-SLEYDLFGNTFEO-PKVLPIVKLQAPITL-LESSARTIHNRI 260
DB 269 TLPIEIGQLQNLQTLNLSNQLTTLTI-EIGKQLQNLQTLNLSNQLTTLTKSI 320

```

RESULT 12	PRELIMINARY;	PRT;	679 AA.
061967			
ID	061967		
AC	061967; Q9NFN7;		
DT	01-AUG-1998 (TrEMBLrel. 07, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Hypothetical protein (LET-413 protein).		
GN	F26D11.11 OR LET-413.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;		
OC	Rhabditidae; Peloderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	MEDLINE=99069613; PubMed=9851916;		
RA	None;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium."		
RL	Science 282:2012-2018(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	Geisel C., Bradshaw H.;		
RT	"The sequence of C. elegans cosmid F26D11.1";		
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	Waterston R.;		
RT	"Direct Submission."		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	Legouis R., Sookhareea S., Boshier J.M., Baillie D.L., Labouesse M.;		
RA	"LET-413 is a basolateral protein required for the assembly of		
RT	adherens junctions in C. elegans."		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.		
CC	- !- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.		
EMBL	AF068716; AAC17752.2; -		
EMBL	AF276590; CAB91651.1; -		
DR	P1R; T33295; T33295.		
DR	WormPep; F26D11.11; CE29778.		
DR	GO; GO:0007242; Pintracellular signaling cascade; IEA.		
DR	InterPro; IPR001611; LRR.		
DR	InterPro; IPR003591; LRR typ.		
DR	InterPro; IPR001478; PDZ.		
DR	Pfam; PF00560; LRR; 11.		
DR	Pfam; PF00595; PDZ; 1.		
DR	SMART; SM00228; PDZ; 1.		
DR	PROSITE; PS0106; PDZ; 1.		
DR	Hypothetical protein.		
SK	SEQUENCE 679 AA; 75335 MW; 60888EDF5AC0F947 CRC64;		
SW			
Query Match	12.5%;	Score 218; DB 5;	Length 679;
Best Local Similarity	39.4%;	Pred. No. 3.5e-10;	
Matches	61; Conservative	17; Mismatches	69; Indels 8; Gaps 3;
QY	80	LEHLSYGLVVRVDMRLCLSKRLDLSHNHKKLPATIGLIHLQELNLDNHNLESF	139
DB	38	LEDNLMTANIKELDRHLSRLHRLDVSDELAVLPAEIGNLTQILNLRN----	92
QY	140	SVALCHSTIQ--KSLRSLSLKNKIKALPVQFCOLQELNKLDDNELIQFCKIGQLIN	197
DB	93	STAKLPDTMNCNKLATLNLSSNPFTRLPETICCSITILSLNETSILTPSNIGSLTN	152
QY	198	LRFLSARKKLPLPSEFNNL-SLEYLDLFGNTFE-231	
DB	153	LRVLEARDNLLRTIPLISVELRKLEELDLDGQNELE 187	

DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR; 5.
DR SMART: SM00366; LRR_PS; 6.
DR SMART: SM00369; LRR_TYP; 5.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 289 AA; 32580 MW; DF38A259B488462C CRC64;

Query Match 12.3%; Score 215; DB 13; Length 289;
Best Local Similarity 27.9%; Pred. No. 2.3e-10;
Matches 75; Conservative 34; Mismatches 104; Indels 56; Gaps 6;
QY 106 LQLSHNHKKLPATIGDLIHQLNELNDNHLESFVALCHSTLQKSLRSLDKNKKIKAL 165
DQ 1 LDISNNNVVIGEILGLTKTLAKNNRLDEFSFKMGGMAR--LEVNLSGNRFEEI 58
QY 166 PVQFCQLQELKNKLDNDNELIQPCIKGQLINRFLSAARKLPFLPSEFRL----- 218
DQ 59 PDQFLQIPTKLSLGGNRLKSPAEIENLISLEFLYLGNGFTISSPSELANLPYLSLV 118
QY 219 -----SLEYLDLFCN--TFEQPKVLPVILKQ----- 242
DQ 119 LCDNRIOSTPPQLAQVHSLSRSLSLNLYLPREILSVHLHLSLRGNFLVVRFDL 178
QY 243 --APUTLESSARTILHNRIPYGSIIIPFHLQDLDTAKICVCGREFCLNSFIQGTITMNL 300
DQ 179 TVTPEPTLEAGTKIKSHGIPYCPWELPENLLRYLDLASKCP-----NPKCSG---VVF 229
QY 301 HSAVHTVVLVDNLGTEAPLISYFCSLGC 329
DQ 230 DCCVRQIKFVDFCGKYRLPLMHYLCSPC 258

RESULT 15
Q96NW7 * PRELIMINARY; PRT; 1537 AA.
AC Q96NW7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Densin-180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Izawa I., Nishizawa M., Ohtakara K., Inagaki M.;
RT "Densin-180 interacts with Delta-Catenin/Neural Plakophilin-Related
Armadillo-Repeat Protein at Synapses.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL: AF434715; AAL28133.1; -.
DR GO: GO:0007242; 2:intracellular signaling cascade; IEA.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003591; LRR typ.
DR InterPro: IPR001478; PDZ.
DR Pfam: PF00560; LRR; 9.
DR Pfam: PF00595; PDZ; 1.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS0106; PDZ; 1.
SQ SEQUENCE 1537 AA; 172580 MW; 1EB99AF133CD2925 CRC64;

Query Match 12.2%; Score 212.5; DB 4; Length 1537;
Best Local Similarity 34.7%; Pred. No. 2.8e-09;
Matches 59; Conservative 27; Mismatches 75; Indels 9; Gaps 4;
QY 75 NPPYSLEHLOTYSGLVRVDMKCLSKRKLDLSHNHKKLPATIGDLIHQLNELNDN 134
DQ 43 NFERTLELYLDANGIEELPKQLFNCQALRKLSIPDNLNLPTTASIVNLKELDISKN 102
QY 135 HLESP--SVALCHSTLQKSLRSLDKNKKIKALPVQFCQLQELKNKLDNDNELIQPCKI 192

Search completed: June 15, 2004, 10:02:46
Job time : 46 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:55:20 ; Search time 17 Seconds

(without alignments)
1035.277 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYLSPVEFIILQLLFIQAI.....PIISYFCSLGCYVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1660	95.1	414	1 LRL1 HUMAN	Q96150 homo sapien
2	201.5	11.5	1256	1 FLR1 DROME	Q24020 drosophila
3	198	11.3	793	1 CDC1 HUMAN	Q9Y219 homo sapien
4	197.5	11.3	810	1 LRC8 HUMAN	Q81w16 homo sapien
5	194	11.1	810	1 LRC8 MOUSE	Q80w95 mus musculus
6	192.5	11.0	2026	1 CYAA YEAST	P08678 saccharomyc
7	183.5	10.5	276	1 RSU1 MOUSE	Q01730 mus musculus
8	183	10.5	1257	1 FLH1 CAEL	P34268 caenorhabdi
9	180.5	10.3	276	1 RSU1 HUMAN	Q15404 homo sapien
10	178.5	10.2	582	1 SHO2 HUMAN	Q9UQ13 homo sapien
11	177.5	10.2	582	1 SHO2 MOUSE	Q88520 mus musculus
12	175	10.0	1839	1 CYAA SACKL	P23466 saccharomyc
13	170	9.7	1039	1 YR71 CAEL	Q09564 caenorhabdi
14	169	9.7	371	1 LRC2 MOUSE	Q8VDB8 mus musculus
15	168	9.6	371	1 LRC2 HUMAN	Q9BYS8 homo sapien
16	165.5	9.5	2493	1 CYAA DUTMA	P49606 ustilago ma
17	165	9.5	1269	1 FLH1 HUMAN	Q13045 homo sapien
18	164	9.4	2300	1 CYAA NEUCR	Q01631 neurospora
19	163.5	9.4	2145	1 CYAA PODAN	Q01513 podospora a
20	158	9.1	423	1 OMD MOUSE	Q35103 mus musculus
21	158	9.1	581	1 LRL5 HUMAN	Q8TF66 homo sapien
22	154	8.8	951	1 LGR4 HUMAN	Q9BXB1 homo sapien
23	154	8.8	1024	1 PCPC RALSO	Q9BBS2 raistonia s
24	148.5	8.5	422	1 OMD BOVIN	Q77742 bos taurus
25	148.5	8.5	560	1 GPV HUMAN	P40197 homo sapien
26	147	8.4	907	1 LGS5 MOUSE	Q921D4 mus musculus
27	146.5	8.4	421	1 OMD HUMAN	Q99983 homo sapien
28	146.5	8.4	662	1 GARP HUMAN	Q14392 homo sapien
29	146	8.4	828	1 LGS6 HUMAN	Q9BXB8 homo sapien
30	146	8.4	1054	1 R131 ARATH	Q91R14 arabidopsis
31	145	8.3	423	1 OMD RAT	Q921S7 rattus norv
32	144.5	8.3	1692	1 CYAA SCHPO	P14605 schizosach
33	144	8.3	343	1 LUM_CHICK	P51890 gallus gall

34	144	8.3	343	1 LUM_COTJA	Q9de67 coturnix co
35	142.5	8.2	1192	1 EXS ARATH	Q91yn8 arabidopsis
36	141.5	8.1	578	1 LRL5 RAT	Q8r5m3 rattus norv
37	141	8.1	331	1 PLI8 AGKBL	Q93233 agkistrodon
38	141	8.1	951	1 LGR4 RAT	Q9Z2H4 rattus norv
39	139	8.0	338	1 LUM_HUMAN	P51884 homo sapien
40	139	8.0	1207	1 BRI1 LYCES	Q8guq5 lycopersico
41	139	8.0	1895	1 WR19 ARATH	Q9Z67 arabidopsis
42	138.5	7.9	342	1 LUM_BOVIN	Q05443 bos taurus
43	138.5	7.9	907	1 LGR5 HUMAN	Q75473 homo sapien
44	138	7.9	782	1 CHAO TRICA	P82963 tribolium c
45	138	7.9	999	1 RLK5 ARATH	Z47735 arabidopsis

ALIGNMENTS

RESULT 1
LRL1_HUMAN
ID LRL1_HUMAN STANDARD; PRT; 414 AA.
AC Q96150; Q86SZ1; Q8N6H9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE LRR-repeat protein 1 (LRR-1) (4-LBB-mediated signaling molecule);
DE (4-1BBLrr)
GN PP1L5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=21662677; PubMed=11804328;
RA Jang I.-K., Lee Z.-H., Kim H.-H., Hill J.M., Kim J.-D., Kwon B.S.;
RT "A novel leucine-rich repeat protein (LRR-1): potential involvement in
4-LBB-mediated signal transduction.";
RL Mol. Cells 12:304-312(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cervical carcinoma;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=B-cell;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallah S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Bulayk S.W.,
Villalon D.K., Muzny D.M., Sogergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Maman A., Rodrigues S., Sanchez A.,
Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley A.C., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May negatively regulate the 4-LBB-mediated signaling
cascades which result in the activation of NK-kappaB and JNK1.
CC -!- SUBUNIT: Interacts with the cytoplasmic domain of TNFRSF9.
CC -!- ALTERNATIVE PRODUCTS:

```

CC CC Event=Alternative splicing; Named isoforms=2;
CC CC Name=1; Synonyms=LRR-1a;
CC CC IsoId=Q96L50-1; Sequence=Displayed;
CC CC Name=2; Synonyms=LRR-1b;
CC CC IsoId=Q96L50-2; Sequence=VSP_008363; VSP_008364;
CC CC -!- TISSUE SPECIFICITY: Ubiquitous. Maximal expression was seen in
CC CC the heart and skeletal muscle and minimal expression seen in
CC CC the kidney.
CC CC -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC CC
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; AY052405; AAL11430.1; -
CC CC EMBL; BX248298; CAD62625.1; ALT_INIT.
CC CC EMBL; BC030142; AAH30142.1; -
CC CC Genew; HGNC:19742; PRL5.
CC CC InterPro; IPR001611; LRR.
CC CC Pfam; PF00560; LRR; 4.
CC CC PRINTS; PRO0019; LEUCRCHPT.
CC CC Repeat; Leucine-rich repeat; Alternative splicing.
CC CC REPEAT 176 199
CC CC REPEAT 201 222
CC CC REPEAT 225 248
CC CC REPEAT 250 271
CC CC REPEAT 273 293
CC CC VARSPLIC 95 146
CC CC
CC CC TMVITSKKDY -> DSIWLSVHSIPSLPRFGYRKNLC:MK
CC CC ILSFHSRNYHESAPCCPHCGLSR (in isoform
CC CC 2).
CC CC
CC CC VARSPLIC 147 414
CC CC Missing (in isoform 2).
CC CC /FTId=VSP_008364.
CC CC G -> V (IN REF. 1).
CC CC A -> S (IN REF. 1).
CC CC
CC CC CONFLICT 25 28
CC CC REPEAT 273 293
CC CC SEQUENCE 414 AA: 46722 MW: 819817BD6CB33C4C CRC64;
CC CC
CC CC Query Match 95.1%; Score 1660; DB 1; Length 414;
CC CC Best Local Similarity 99.7%; Pred. No. 1.2e-120;
CC CC Matches 320; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC CC
QY 18 QAISSLKGFSLNRLAHRGCVNDTPVSTLTPVKTSEFENFKTKWITSKKYPLSKNFP 77
Db .:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 KAISSLKGFSLNRLAHRGCVNDTPVSTLTPVKTSEFENFKTKWITSKKYPLSKNFP 153
QY 78 YSLEHLQTSYCGLVYRDMRLCLKSLRKLDSLNNHKKLPATIGDLIHQELMNDNHL 137
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 YSLEHLQTSYCGLVYRDMRLCLKSLRKLDSLNNHKKLPATIGDLIHQELMNDNHL 213
QY 138 SFSVALCHSTLQKSLRSLDLKSKNKIKALPVQFCQLQELKNLKDDELIOFPCKIGLIN 197
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 SFSVALCHSTLQKSLRSLDLKSKNKIKALPVQFCQLQELKNLKDDELIOFPCKIGLIN 273
QY 198 LRFSLAARNKLPLPSEFNLSLEYLDLFGNTFQPKVLPVKLQAPLTLLESSARTILH 257
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 LRFSLAARNKLPLPSEFNLSLEYLDLFGNTFQPKVLPVKLQAPLTLLESSARTILH 333
QY 258 NRIPYGSHPHFLCQDLDTAKICVGRFCNLSFIQGTITMNLHLSVAHTVVLVDNLGGTE 317
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 NRIPYGSHPHFLCQDLDTAKICVGRFCNLSFIQGTITMNLHLSVAHTVVLVDNLGGTE 393
QY 318 APIISYFCSIGCVYNSDMLK 338
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
394 APIISYFCSIGCVYNSDMLK 414
CC CC RESULT 2
CC CC FLII_DRONE

```

```

ID ID FLII_DRONE STANDARD; PRT; 1256 AA.
AC Q24020; Q24088; Q9VRH0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Flightless-I protein.
GN FLII OR CGI484.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=94068608; PubMed=8248253;
RA Campbell H.D., Schimanski T., Claudianos C., Ozsarac N.,
RA Kasprzak A.B., Coatsell C.N., Young I.G., de Couet H.G., Miklos G.L.G.;
RT "The Drosophila melanogaster flightless-I gene involved in
RT gastrulation and muscle degeneration encodes gelsolin-like and
RT leucine-rich repeat domains and is conserved in Caenorhabditis elegans
RT and humans.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390 (1993).
RN [2]
SEQUENCE FROM N.A., AND VARIANT SER-601.
RC STRAIN=Canton-S;
RX MEDLINE=96129280; PubMed=8582612;
RA de Couet H.G., Fong K.S.K., Weeds A.G., McLaughlin P.J.,
RA Miklos G.L.G.;
RT "Molecular and mutational analysis of a gelsolin-family member encoded
RT by the flightless I gene of Drosophila melanogaster.";
RL Genetics 141:1049-1053 (1995).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98188272; PubMed=9520435;
RA Maleszka R., de Couet H.G., Miklos G.L.G.;
RT "Data transferability from model organisms to human beings: insights
RT from the functional genomics of the flightless region of Drosophila.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:3731-3736 (1998).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=20136012; PubMed=10731138;
RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Fzise E.,
RA Stapleton M., Harvey D.A.;
RT "A Drosophila complementary DNA resource.";
RL Science 287:2222-2224 (2000).
RN [5]
SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20136006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Jahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB023233; BAA76860.1; --
 Genew; HGNC:120309; CHDC1.
 InterPro; IPR001715; Calponin-like.
 InterPro; IPR001611; LRR.
 InterPro; IPR003591; LRR_type.
 Pfam; PF00307; CH; 1.
 Pfam; PF00560; LRR; 6.
 SMART; SM00033; CH; 1.
 SMART; SM00369; LRR_type; 1.
 PROSITE; PS00021; CH; 1.
 Leucine-rich repeat; Repeat.
 NON_TER 1
 DOMAIN 86 101 HIS-RICH.
 DOMAIN 102 119 GLY-RICH.
 REPEAT 135 158
 REPEAT 161 184
 REPEAT 185 207
 REPEAT 208 230
 REPEAT 232 252
 REPEAT 233 275
 REPEAT 276 298
 REPEAT 300 319
 REPEAT 320 343
 DOMAIN 641 754 CH.
 SEQUENCE 793 AA; 87801 MW; 0BE99CC48CBB37C CRC64;
 Query Match 11.3%; Score 198; DB 1; Length 793;
 Best Local Similarity 34.8%; Pred. No. 1e-07;
 Matches 62; Conservative 22; Mismatches 70; Indels 24; Gaps 3;
 QY 72 LSKNFPYSLEHLQTSYGLVRYDVMRLCLKSLRKLDSLNHHIKLPATIGDLHLOSLNL 131
 Db 159 LSKN-----RLVRFMELCHPVSLVLELNLYNCIRVIPPEAVINQLMITYLNL 215
 QY 132 NDNHLESFVALCHSTLQSLRSLDSLNKIKALPVQFQLOELKNLKDNDLELQFPCK 191
 Db 216 SRNQLSALPACLC---GLPLKVLASNNKLSLPEEIGQLQMLDVSCHNEITLPPQ 271
 QY 192 IGLINLRFSLAARNLPLPSEFNLSLEYLDLFGN-----TFEPQKVLPIVKLQ 242
 Db 272 IGLKSLRELNVRRNYLKVLPQELVDLPVDFDSCNKKVLVPIPCFRENKQLQVLLLE 329

RESULT 4

LRC8 HUMAN
 ID LRC8 HUMAN STANDARD; PRT; 810 AA.
 AC Q81WT6; Q8NC10; Q9P2B1;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Leucine-rich repeat-containing protein 8 precursor.
 GN LRC8 OR KIAA1437.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sawada A., Takihara Y., Kim J.-Y., Matsuda-Hashii Y., Tokimasa S.,
 RA Fujisaki H., Kubota K., Endo H., Onodera T., Ohta H., Hara J.;
 RA "A novel gene LRR8 is essential to the B cell development.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirokawa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 The complete sequences of 150 new cDNA clones from brain which code

RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Young A.C., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 259-810 FROM N.A.
 RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Makamura Y., Kojima S., Nagahara K.,
 RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AV143166; RAN18279.1; --

EMBL; AB037858; BAA92675.1; ALT_INIT.

EMBL; BC051322; BAA51322.1; --

EMBL; AK074723; BAC11161.1; ALT_INIT.

Genew; HGNC:19027; LRR8.

InterPro; IPR001611; LRR.

InterPro; IPR003591; LRR_type.

Pfam; PF00560; LRR; 7.

PRINTS; PR00019; LEURICHRPT.

Repeat; Leucine-rich repeat; Signal.

SIGNAL 1 810

POTENTIAL.

LEUCINE-RICH REPEAT-CONTAINING PROTEIN 8.

LRR 1.

LRR 2.

LRR 3.

LRR 4.

LRR 5.

LRR 6.

LRR 7.

LRR 8.

LRR 9.

LRR 10.

LRR 11.

LRR 12.

LRR 13.

LRR 14.

GLCNAC. . . (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

83

CC CARBOHYD 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 554 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 592 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 660 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 755 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 755 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 323 S -> P (IN REF. 4).
 CC CONFLICT 744 Q -> R (IN REF. 4).
 CC SEQUENCE 810 AA; 94198 MW; E0C5E6BEE8275E8 CRC64;

Query Match 11.3%; Score 197.5; DB 1; Length 810;
 Best Local Similarity 29.3%; Pred. No. 1.2e-07;
 Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTKWITSKKDYPLSKNFPYSLEHLOTSYGLVVRMRLCLKSLKRLKLS----- 109
 DB 576 NEGTKLIVLN-----SLKKVNLTELELRCDLERIPHSIFSLHNLQELDKNNLXTI 629
 QY 110 -----HNHIKKLPATIGDILHLOELNLDNHLNLSFVALCHSTLOKS 151
 DB 630 EELISFQHLRLTCLKLNWNIAYIPIQIGNLTNLERLYLNENKIEKIPQLFYC---RK 686
 QY 152 LRSLDLNNKIKALPVQFCQQLKLNLDLDDNELIOPCKIQLINLRLSARNKLPPL 211
 DB 687 LRYLDSHNNLTFLPADIGLLQNLQNLAVTANRIALPPELPFCCKLRALHGNVQLSL 746
 QY 212 PSEFENL-SLEYLDLFGNTFEQPKVLPVTKIQAPLTLLESSARTI---LHNRIP 261
 DB 747 PSRVGELNLQTLRLGRNLE---CLPVELGECPL-LKRSGLVVEEDLFNLTLP 795

RESULT 5
 ID LCRC MOUSE STANDARD; PRT; 810 AA.
 AC Q80WG5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Leucine-rich repeat-containing protein 8 precursor.
 GN LRC8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Klausner R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millhys S.J.,
 RA Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Skalski D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
 CC This SWISS-PROF entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; BC048152; AAR48152.1; -;
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR003885; LRR_CYS.
 CC InterPro; IPR003591; LRR_Typ.
 CC 2fam; PF00560; LRR; 6;
 CC PRINTS; PR00019; LEURICHRPT.
 CC SMART; SM00366; LRR_PS; 8.
 CC SMART; SM00365; LRR_SD22; 3.
 CC SMART; SM00369; LRR_TVP; 8.
 KW Repeat; Leucine-rich repeat; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN ? 810
 FT REPEAT 38 62
 FT REPEAT 422 445
 FT REPEAT 492 515
 FT REPEAT 518 542
 FT REPEAT 543 565
 FT REPEAT 590 612
 FT REPEAT 613 637
 FT REPEAT 639 660
 FT REPEAT 661 684
 FT REPEAT 706 730
 FT REPEAT 732 752
 FT REPEAT 753 776
 FT REPEAT 778 801
 FT CARBOHYD 66 66
 FT CARBOHYD 83 83
 FT CARBOHYD 191 191
 FT CARBOHYD 526 526
 FT CARBOHYD 554 554
 FT CARBOHYD 592 592
 FT CARBOHYD 660 660
 FT CARBOHYD 696 696
 FT CARBOHYD 755 755
 FT SEQUENCE 810 AA; 94119 MW; E3B125BC9977DBB0 CRC64;
 Query Match 11.1%; Score 194; DB 1; Length 810;
 Best Local Similarity 29.3%; Pred. No. 2.1e-07;
 Matches 63; Conservative 30; Mismatches 84; Indels 38; Gaps 5;

QY 57 NFKTKWITSKKDYPLSKNFPYSLEHLOTSYGLVVRMRLCLKSLKRLKLS----- 109
 DB 576 NEGTKLIVLN-----SLKKVNLTELELRCDLERIPHSIFSLHNLQELDKNNLXTI 629
 QY 110 -----HNHIKKLPATIGDILHLOELNLDNHLNLSFVALCHSTLOKS 151
 DB 630 EELISFQHLRLTCLKLNWNIAYIPIQIGNLTNLERLYLNENKIEKIPQLFYC---RK 686
 QY 152 LRSLDLNNKIKALPVQFCQQLKLNLDLDDNELIOPCKIQLINLRLSARNKLPPL 211
 DB 687 LRYLDSHNNLTFLPADIGLLQNLQNLAVTANRIALPPELPFCCKLRALHGNVQLSL 746
 QY 212 PSEFENL-SLEYLDLFGNTFEQPKVLPVTKIQAPL 245
 DB 747 PSRVGELNLQTLRLGRNLE---CLPVELGECPL 778

RESULT 6
 CYAA YEAST
 ID CYAA YEAST STANDARD; PRT; 2026 AA.
 AC P08678;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
 DE cyclase).

GN CYR1 OR CDC35 OR HSR1 OR SRA4 OR YOL005W OR J1401.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86079531; PubMed=2934138;
RA Kataoka T., Broek D., Wigler M.;
RT "DNA sequence and characterization of the S. cerevisiae gene encoding
RL adenylyl cyclase.";
RN Cell 43:493-505 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1042-2026 FROM N.A.
RX MEDLINE=88165073; PubMed=3327602;
RA Masson P., Lenzén G., Jacquemin J.M., Danchin A.;
RT "Yeast adenylyl cyclase catalytic domain is carboxy terminal.";
RL Curr. Genet. 10:343-352 (1986).
RN [5]
RP MUTAGENESIS OF THR-1651.
RX MEDLINE=91122042; PubMed=191451;
RA Feger G., de Vendittis E., Vitelli A., Masturzo P., Zahn R.,
RA Verrotti A.C., Kavounis C., Pal G.P., Fasano O.;
RT "Identification of regulatory residues of the yeast adenylyl
RL cyclase.";
RN EMBO J. 10:349-359 (1991).
CC -!- FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC cAMP.
CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -!- ENZYME REGULATION: The presence of GTP-bound Ras2 protein is
CC required in order to elicit a magnesium-dependent adenylyl cyclase
CC activity.
CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 PP2C-like domain.
CC -!- SIMILARITY: Contains 1 Ras-associating domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12057; AAA34549.1; -;
CC EMBL; X49283; CAA89295.1; -;
CC EMBL; X87611; CAA60917.1; -;
CC EMBL; X03449; CAA27175.1; -;
CC PIR; S56776; OVB;
CC GeneOnline; 141621; -;
CC SGD; S0003542; CYR1.
CC GO; GO:0005886; C:plasma membrane; IDA.
CC GO; GO:0007126; P:meiosis; IMP.
CC GO; GO:0000074; P:regulation of cell cycle; IMP.
CC InterPro; IPR001054; G: cyclase.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR_type.
CC InterPro; IPR001932; PP2C-like.
CC InterPro; IPR000159; RA domain.
CC Pfam; PF00211; Guanylate_cyc; 1.
CC Pfam; PF00560; LRR; 15.

DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00044; CYC; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00314; RA; 1.
DR PROSITE; PSS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PSS0200; RA; 1.
KW Lyase; Repeat; Leucine-rich repeat; CAMP biosynthesis; Metal-binding;
KW Magnesium.
KW DOMAIN 676 755 RAS-ASSOCIATING.
FT REPEAT 815 838 LRR 1.
FT REPEAT 842 862 LRR 2.
FT REPEAT 863 885 LRR 3.
FT REPEAT 886 908 LRR 4.
FT REPEAT 910 931 LRR 5.
FT REPEAT 932 955 LRR 6.
FT REPEAT 957 976 LRR 7.
FT REPEAT 977 999 LRR 8.
FT REPEAT 1001 1016 LRR 9.
FT REPEAT 1017 1040 LRR 10.
FT REPEAT 1042 1062 LRR 11.
FT REPEAT 1063 1086 LRR 12.
FT REPEAT 1088 1109 LRR 13.
FT REPEAT 1110 1132 LRR 14.
FT REPEAT 1134 1156 LRR 15.
FT REPEAT 1188 1209 LRR 16.
FT REPEAT 1210 1232 LRR 17.
FT REPEAT 1233 1256 LRR 18.
FT REPEAT 1258 1280 LRR 19.
FT REPEAT 1285 1308 LRR 20.
FT REPEAT 1319 1344 LRR 21.
FT DOMAIN 1369 1625 PP2C-LIKE.
FT DOMAIN 1626 2026 CATALYTIC.
FT METAL 1673 1673 MAGNESIUM (BY SIMILARITY).
FT METAL 1716 1716 MAGNESIUM (BY SIMILARITY).
FT MUTAGEN 1651 1651 T->I: WEAK RAS-INDEPENDENT ACTIVITY.
FT MUTAGEN 1651 1651 T->I: ATTENUATION OF THE RESPONSE TO
FT RAS PROTEINS.
FT CONFLICT 262 262 S -> L (IN REF. 1).
FT CONFLICT 548 548 S -> L (IN REF. 1).
FT CONFLICT 592 592 D -> H (IN REF. 1).
FT CONFLICT 709 709 R -> I (IN REF. 1).
FT CONFLICT 962 962 L -> P (IN REF. 1).
FT CONFLICT 1388 1388 L -> S (IN REF. 1).
FT CONFLICT 1427 1427 E -> D (IN REF. 3).
FT CONFLICT 1461 1461 A -> T (IN REF. 3).
FT CONFLICT 1566 1566 A -> S (IN REF. 1).
FT CONFLICT 1735 1735 V -> G (IN REF. 1).
FT CONFLICT 1956 1956 I -> V (IN REF. 3).
FT CONFLICT 1996 1996 C -> F (IN REF. 1).
FT CONFLICT 2009 2026 NVYDELLQWVKNKADLST -> MLLTNFYKMLRTQRIYQL
FT CONFLICT 2026 AA; 227832 MW; BEAB7419AB6989D0 CRC64;
SQ SEQUENCE 2026 AA; 227832 MW; BEAB7419AB6989D0 CRC64;
Query Match 11.0%; Score 192.5; DB 1; Length 2026;
Best local similarity 25.9%; Pred. No. 8.4e-07;
Matches 84; Conservative 52; Mismatches 125; Indels 63; Gaps 13;
QY 2 YVLSVPEFIIQLLFTQALSSSLKGLSAMLAHRCNVDPVSTLTGVKTSFENFKTK 61
DB 863 YKLVSLE---LQNFIRKVPNSIMK-LSNLTINLQCN---ELESF-PAGFVLEKQLQL 914
QY 62 MVTSK-KDYPLSKNPPYSLEHLQTSYGLVRVDMMLCLSKRLKDLGHNHKKLPAT 120
DB 915 DLSSNKFMYPEVINYCTNLQIDLQSLQSLQSLQSLQSLQSLQSLQSLQSLQSL 120
QY 121 GDL-----IHLELNINHNHLESFVAL----- 143
DB 971 GDLSEMTDLRTLNLRYNRISSIKTNASNLQNLFLTDNRISNEEDTLPKLRALFIOENPIT 1030
QY 144 ---CHSTLQKSLRSILDLKSKKIKALPVQ-FCOLOELKLNKLDNDELIOFPCKIGQLINLR 199
DB 1031 SISPKDFYFKNWTSLTNKKAQLSSIPGELLTKLSFLKLELNQNLTRLPQELSKTLV 1090

QY 230 FLSARNKLPFLPSPERNL-SLEYDLFGNTP-----EQPKVLPVVKIQAPLTL-ESSAR 253
 DB 1091 FLSVARNKLEYIPPELSQKSLRTLDLHNHNRDQVDFGKMNLELTSLNSSNAFGNSSLE 1150
 QY 254 TLHNRIKPVGSHI---IPEHLQCD 274
 DB 1151 NSFYHNSYSGSKLSKSLMFFIAAD 1174

RESULT 7
 RSUL_MOUSE
 ID RSUL_MOUSE STANDARD; PRT; 276 AA.
 AC Q01730;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ras suppressor protein 1 (Rsu-1) (RSP-1).
 GN RSU1 OR RSP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH Swiss; TISSUE=Fibroblast;
 RX MEDLINE=92375042; PubMed=1508180;
 RA Cutler M.E., Bassin R.H., Zannoni L., Talbot N.;
 RT "Isolation of rsp-1, a novel cDNA capable of suppressing v-Ras
 transformation.";
 RL Mol. Cell. Biol. 12:3750-3756 (1992).
 CC -!- FUNCTION: Potentially plays a role in the Ras signal transduction
 pathway. Capable of suppressing v-Ras transformation in vitro.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X63039; G244765.1; -
 DR PIR; S25770; S25770.
 DR MGI; MGI:133040; Rsul.
 DR GO; GO:0037265; P-RAS protein signal transduction; IMP.
 DR InterPro; IPR001611; LRR.
 DR Pfam; PF00560; LRR; 7
 DR SMART; SM00364; LRR_BAC; 5.
 DR SMART; SM00366; LRR_PS; 6.
 DR SMART; SM00369; LRR_TYP; 6.
 DR Leucine-rich repeat; Repeat.
 DR INIT_MET 0 BY SIMILARITY.
 FT REPEAT 38 61 LRR 1.
 FT REPEAT 63 84 LRR 2.
 FT REPEAT 85 107 LRR 3.
 FT REPEAT 109 132 LRR 4.
 FT REPEAT 133 155 LRR 5.
 FT REPEAT 157 178 LRR 6.
 FT REPEAT 179 201 LRR 7.
 SQ SEQUENCE 276 AA; 31419 MW; 245FE4AA27AF16DE CRC64;

Query Match 10.5%; Score 183.5; DB 1; Length 276;
 Best Local Similarity 31.2%; Pred. No. 3.8e-07;
 Matches 60; Conservative 26; Mismatches 81; Indels 25; Gaps 3;
 QY 72 LSKNFPYSLEH---CTSYCYGLVRVDMCLCKSLRKLDSLNHNKKKLPATIGDIHQE 128
 DB 30 LDVNGDLSLAHTQTQLSHNKLTTVPNVAVLKNLELVNFFNNQIEELTQISSIQKLKH 89
 QY 129 LNLNDNHEGFS-----VALCHSTIQK-----SLRSLDSKNKIKALP 166

DB 90 INLGMNRLNTLPRGFGSSRLLEVLTYMNLNHSHPGNFFYLTTRALYLSNDFFILP 149
 QY 167 VQFCQELKKNLKDNLQIQLQPKIGQNLNRLFLSAARNKLPFLPSEFRNLSLEYLDLF 226
 DB 150 PDIGKLTQKQLSLRDLNLSLPKELGELTQKELHIQGNRLVLPPELGNLDLTGQXQV 209
 QY 227 GNTFEQPKVLPV 238
 DB 210 FKAENNEWVTEI 221

RESULT 8
 FLIH_CABEL
 ID FLIH_CABEL STANDARD; PRT; 1257 AA.
 AC P34268; P90739; Q1731;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flightless-I protein homolog.
 GN B0523.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94068608; PubMed=8248259;
 RX Campbell H.D., Schimansky T., Claudianos C., Ozsarac N.,
 RA Kagrzak A.B., Cotsell J.N., Young I.G., de Couet H.G., Miklos G.L.G.;
 RT "The Drosophila melanogaster flightless-I gene involved in
 RT gastrulation and muscle degeneration encodes gelsolin-like and
 RT leucine-rich repeat domains and is conserved in Caenorhabditis
 RT elegans and humans.";
 RN Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Alpcough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulston L., Gardner A., Green P., Hawkins T., Hillier L., Jier X.,
 RA Johnston L., Jones M., Kerhaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38 (1994).
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (JUN-1996) to Swiss-Prot.
 CC -!- FUNCTION: May play a key role in embryonic cellularization by
 CC interacting with both the cytoskeleton and other cellular
 CC components (by similarity).
 CC -!- SIMILARITY: Belongs to the villin/gelsolin family.
 CC -!- SIMILARITY: Contains 16 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 4 gelsolin-like repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U01183; AAC03567.1; -
 DR EMBL; L07143; AAB37088.1; -

```
DR PIR; A88536; A88536.
DR PIR; S27783; S27783.
DR HSSP; P10733; 1SVY.
DR WormPep; B0523.5; CE06721.
DR InterPro; IPR007122; Gelsolin.
DR InterPro; IPR007123; Gelsolin.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00626; Gelsolin; 4.
DR Pfam; PF00560; LRR; 10.
DR PRINTS; PR00597; GELSOLIN.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SM00262; GEL; 6.
DR SMART; SM00369; LRR TYP; 1.
DR Developmental protein; Repeat; Leucine-rich repeat.
FT REPEAT 6 31 LRR 1.
FT REPEAT 32 54 LRR 2.
FT REPEAT 56 77 LRR 3.
FT REPEAT 78 102 LRR 4.
FT REPEAT 103 126 LRR 5.
FT REPEAT 128 148 LRR 6.
FT REPEAT 149 172 LRR 7.
FT REPEAT 174 195 LRR 8.
FT REPEAT 197 221 LRR 9.
FT REPEAT 222 244 LRR 10.
FT REPEAT 246 267 LRR 11.
FT REPEAT 268 290 LRR 12.
FT REPEAT 292 315 LRR 13.
FT REPEAT 316 338 LRR 14.
FT REPEAT 339 361 LRR 15.
FT REPEAT 363 384 LRR 16.
FT REPEAT 510 568 GELSOLIN-LIKE 1.
FT REPEAT 641 681 GELSOLIN-LIKE 2.
FT REPEAT 755 798 GELSOLIN-LIKE 3.
FT REPEAT 1166 1206 GELSOLIN-LIKE 4.
FT CONFLICT 154 154 P -> S (IN REF. 3).
SQ SEQUENCE 1257 AA; 144814 MW; F3FED79FACB07493 CRC64;

Query Match
Best Local Similarity 10.5%; Score 183; DB 1; Length 1257;
Matches 81; Conservative 46; Mismatches 101; Indels 100; Gaps 12;

QY 14 LFIQIAISSLKGFLSAMLAHRCNCVDPVSLTPVKTSSEFNFKWVITKQYPLS 73
Db 6 LQFVGIDFSGNDF-SGDRFPH-----DVEQMTQMTWLKLNDSKLEQV-----PDE 50
QY 74 KNFPVSLHLOTYSYGLVRYDMCLCKSLRK-----LDL 108
Db 51 LSRCANLEHQAHHNQLISVHGELSDLPRLRSVIVRDNNLKTAGIPTDIFPMKDLTIIDL 110
QY 109 SHNHKKLPATI-----GDLIHLQELMNDHLESFVALCHSTLQSLRSLDLKSKNIKA 164
Db 111 SRNQLREVPTNLEYAGSIV---LNLSYNIETIPNSVCANLI--DLLFLDLNNKLDLM 164
QY 165 LPVQFCQLQELKXKLDDNELIQF-----PCKIGQLINL 198
Db 165 LPPQRLRLSMLOSRLKSNPLNEFQKLPQSMFSLSVLHMSNTNRTLDNIPPTLDMDNML 224
QY 199 RFLSARNKLPFLPSR-FRNLSLEYLDLFCNTPE-----QPKVLP 237
Db 225 RDVDFSENNLPVPEALFKLRNKLNLKNSKLEKLNMTGEWENLETNMSINQITVLP 284
QY 238 --VIKIQAPLTLESARTILHNIRPYG 263
Db 285 DCVVKL-TRLTLYAANNQITPEGIFSG 311

RESULT 9
RSUL_HUMAN
ID RSUL_HUMAN STANDARD; PRT; 276 AA.
AC Q15404;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
```

RT	RL	RN	for DNA [4]
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

```
Query Match      10.2%; Score 178.5; DB 1; Length 582;
Best local Similarity 30.8%; Pred. No. 2.3e-06;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLRKLDSHHHKKLPATIGDLHLQELNLDNHLSFVSVALCHSTLQKSLRSLDLSKNK 161
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 403 SWVELNLTNQTKIPEDVSGVLSLEVLILSNLLKKLPGL--GNLRK-LRELDLEENK 459
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 IKALPVQFCQLOELKLDNDELIOFPCKIGQLINRFLSAARNKLPFLPSFRNL-SL 220
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 LESLPNEIAYLKDQLKVLTNQTLTPRGIGHLTNLTGLGENLLTHLPEEIGTLENL 519
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 EYLDLFGNTFQPKVLPVKLOAPLFLLESSARTILHNRIPIYGSIIIPFH--LCQDLDTA 278
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 520 EELYNDN-----PNE-----HSLPFELALCSKLSIM 546
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 279 KICVC 283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 547 SIENC 551

RESULT 11
SHO2 MOUSE
ID SHO2 MOUSE STANDARD; PRT; 592 AA.
AC O88520;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leucine-rich repeat protein SHOC-2 (Ras-binding protein Sur-8).
GN SHOC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337190; PubMed=9674433;
RA Sieburth D.S., Sun Q., Han M.;
RT "SUR-8, a conserved Ras-binding protein with leucine-rich repeats,
RL Cell 94:119-130(1998).
CC -! SUBUNIT: Specifically binds K-Ras and N-Ras but not H-Ras in
CC vitro.
CC -! SIMILARITY: Contains 20 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF068921; AAC40175.1; -
CC MGD; MGI:1927197; Shoc2.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR_TYP.
CC Pfam; PF00560; LRR; 16.
CC PRINTS; PR00019; LEURICRPT.
CC SMART; SM00369; LRR_TYE; 1.
CC Repeat; Leucine-rich repeat.
FT REPEAT 99 122 LRR 1.
FT REPEAT 123 145 LRR 2.
FT REPEAT 146 168 LRR 3.
FT REPEAT 169 191 LRR 4.
FT REPEAT 193 214 LRR 5.
FT REPEAT 215 237 LRR 6.
FT REPEAT 239 260 LRR 7.
FT REPEAT 262 283 LRR 8.
FT REPEAT 284 306 LRR 9.
FT REPEAT 307 329 LRR 10.
FT REPEAT 331 353 LRR 11.
FT REPEAT 354 377 LRR 12.

Query Match      10.2%; Score 177.5; DB 1; Length 582;
Best local Similarity 30.8%; Pred. No. 2.7e-06;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLRKLDSHHHKKLPATIGDLHLQELNLDNHLSFVSVALCHSTLQKSLRSLDLSKNK 161
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 403 SWVELNLTNQTKIPEDVSGVLSLEVLILSNLLKKLPGL--GNLRK-LRELDLEENK 459
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 IKALPVQFCQLOELKLDNDELIOFPCKIGQLINRFLSAARNKLPFLPSFRNL-SL 220
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 LESLPNEIAYLKDQLKVLTNQTLTPRGIGHLTNLTGLGENLLTHLPEEIGTLENL 519
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 EYLDLFGNTFQPKVLPVKLOAPLFLLESSARTILHNRIPIYGSIIIPFH--LCQDLDTA 278
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 520 EELYNDN-----PNE-----HSLPFELALCSKLSIM 546
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 279 KICVC 283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 547 SIENC 551

RESULT 12
CYAA SACKL
ID CYAA SACKL STANDARD; PRT; 1839 AA.
AC P23466;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenyl-
DE cyclase).
GN CYR1
OS Saccharomyces kluyveri (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9133718; PubMed=1864503;
RA Young D., O'Neill K., Broek D., Wigler M.;
RT "The adenylate cyclase-encoding gene from Saccharomyces kluyveri."
RL Gene 102:129-132(1991).
CC -! FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC CAMP.
CC -! CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -! COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -! SIMILARITY: Belongs to the adenylate cyclase class-3 family.
CC -! SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -! SIMILARITY: Contains 1 PP2C-like domain.
CC -! SIMILARITY: Contains 1 Ras-associating domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56042; CAA39513.1; -
CC InterPro; IPR001054; G_cyclase.
CC InterPro; IPR001611; LRR.
```



```

DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000159; RA domain.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR Pfam; PF00560; LRR; 15.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00044; CYCC; 1.
DR SMART; SM00369; LRR_TYP; 2.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00314; RA; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS0200; RA; 1.
KW Lyase; Repeat; leucine-rich repeat; cAMP biosynthesis; Metal-binding;
KW Magnesium.
FT DOMAIN 494 574 RAS-ASSOCIATING.
FT REPEAT 632 655 LRR 1.
FT REPEAT 659 679 LRR 2.
FT REPEAT 680 702 LRR 3.
FT REPEAT 703 724 LRR 4.
FT REPEAT 726 748 LRR 5.
FT REPEAT 749 771 LRR 6.
FT REPEAT 773 793 LRR 7.
FT REPEAT 794 818 LRR 8.
FT REPEAT 820 833 LRR 9.
FT REPEAT 834 857 LRR 10.
FT REPEAT 859 879 LRR 11.
FT REPEAT 880 903 LRR 12.
FT REPEAT 905 926 LRR 13.
FT REPEAT 927 949 LRR 14.
FT REPEAT 951 974 LRR 15.
FT REPEAT 1004 1025 LRR 16.
FT REPEAT 1026 1048 LRR 17.
FT REPEAT 1050 1072 LRR 18.
FT REPEAT 1074 1096 LRR 19.
FT REPEAT 1101 1124 LRR 20.
FT REPEAT 1135 1160 LRR 21.
FT DOMAIN 1185 1440 PP2C-LIKE.
FT DOMAIN 1441 1839 CATALYTIC.
FT METAL 1488 1488 MAGNESIUM (BY SIMILARITY).
FT METAL 1531 1531 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCB1F273CB CRC64;

Query Match 10.0%; Score 175; DB 1; Length 1839;
Best Local Similarity 26.3%; Pred. No. 1.7e-05;
Matches 69; Conservative 49; Mismatches 88; Indels 56; Gaps 11;

QY 12 LQLLPIQAISSSLKGF-LSAMELAHRCNVDPVSTLTVPKTSSEFENFKTKMVTSSK-- 68
DB LERNFIKKVPDSI--FKLNNLTIVNLCN--NLERLPP-GSKLKNLQL-LDISNKFV 739
QY 69 DYPLSKNFPYSLEHLOTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLI---- 124
DB DYPLSKNFPYSLEHLOTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLI---- 124
QY 740 NYPEVINSCTNLQIDLSYNKTHSLPVSINQLVKLAKNLFNNRL----TSVGDLSQMK 795
DB NYPEVINSCTNLQIDLSYNKTHSLPVSINQLVKLAKNLFNNRL----TSVGDLSQMK 795
QY 125 -----HLQELNLDNHLSEF-----SVALCHSTL 148
DB LRTNLCRNVTSIECHAPNQLNLFNTDRIETFDODLTRLETLQQNPI-SWVGNGY 855
QY 149 QKSRLSLSKNKKIKALPVQ-FCQELKLNKLDNDELITQFCKTGQNLNRLFLSA 207
DB QKSRLSLSKNKKIKALPVQ-FCQELKLNKLDNDELITQFCKTGQNLNRLFLSA 207
QY 856 MANMTSLNSLNKAKLSFSAELLSKLPUEKLENNNTQLPPEINKLTLRLYSVARNK 915
DB MANMTSLNSLNKAKLSFSAELLSKLPUEKLENNNTQLPPEINKLTLRLYSVARNK 915
QY 208 LPFLSEFRNL-SLEYLDLFGN 228
DB LPFLSEFRNL-SLEYLDLFGN 228
QY 916 LESIPDEISDLRLSKSLDLSN 937
DB LESIPDEISDLRLSKSLDLSN 937

RESULT 13
YR71_CABEL STANDARD; PRT; 1039 AA.
ID YR71_CABEL AC Q09564;
AC Q09564;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F43C1.1 in chromosome III.
GN F43C1.1
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Jassal B.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: Contains 18 leucine-rich (LRR) repeats.
CC -|- SIMILARITY: Contains 1 PP2C-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL; Z46937; CAA87056.1; -.
DR PIR; T22117; T22117.
DR WormPep; F43C1.1; CE01582.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF00481; PP2C; 1.
DR PRINTS; PR00319; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
DR SMART; SM00332; PP2C; 1.
KW Hypothetical protein; Leucine-rich repeat; Repeat.
FT REPEAT 137 160 LRR 1.
FT REPEAT 161 185 LRR 2.
FT REPEAT 199 222 LRR 3.
FT REPEAT 223 245 LRR 4.
FT REPEAT 246 268 LRR 5.
FT REPEAT 269 291 LRR 6.
FT REPEAT 293 316 LRR 7.
FT REPEAT 332 358 LRR 8.
FT REPEAT 364 389 LRR 9.
FT REPEAT 396 419 LRR 10.
FT REPEAT 420 442 LRR 11.
FT REPEAT 444 464 LRR 12.
FT REPEAT 465 487 LRR 13.
FT REPEAT 489 514 LRR 14.
FT REPEAT 536 559 LRR 15.
FT REPEAT 561 585 LRR 16.
FT REPEAT 587 605 LRR 17.
FT REPEAT 606 627 LRR 18.
FT DOMAIN 669 903 PP2C-LIKE.
SQ SEQUENCE 1039 AA; 118182 MW; 0C6D1AFC79A0D32A CRC64;

Query Match 9.7%; Score 170; DB 1; Length 1039;
Best Local Similarity 25.3%; Pred. No. 2e-05;
Matches 77; Conservative 52; Mismatches 121; Indels 54; Gaps 11;

QY 24 LKGFISAMELAHRCNVDPVSTLTVPKTSSEFENFKTKMVTSSKDYPLSKNFPYSLEHL 83
DB LEAWLTRAQQVEKSNHVDASDEQLTLPEQILNNEARVQILNLRNLSLRP---PTEKV 183
QY 84 QTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLIHLQELNLDNHLSEFVAL 143
DB QTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLIHLQELNLDNHLSEFVAL 143
DB QKSMAPLGYID-----DLRYVH-----SLQVDSLANSQILSPFIOL 219
QY 144 CHSTLQKSLRLSLSKNKKIKALPVQ-FCQELKLNKLDNDELITQFCKTGQNLNRLFLSA 203
DB CHSTLQKSLRLSLSKNKKIKALPVQ-FCQELKLNKLDNDELITQFCKTGQNLNRLFLSA 203
DB ---TLSSHRLQNLSSNTYSSVPSSECSNMRRLOYNLNNNQLDT-PDSISELQNLQSLDI 276
QY 204 ARNKK-FFLPSEFRNL-SLEYLDLFGNFTFEQ-PKV--LPVTKLQAP-TLLESSAR----TIL 256

```

```

Db 277 SFNQSQIPPCPLPHLTLEMLRAGNIEKIDRGEMQIQKIDLRNVLDTSFRLDIENIT 336
QY 257 HNRIPYGVSHIPLHLCQDIDTAKICVGRFLNSFQIGITIMNLH-----SVAHTVV---L 309
Db 337 HLDLRNSMISTVHL-TNLRFLVKVHCERLQSS-----IHLGSGSLTVEYADHNL 386
QY 310 VDNL 313
Db 387 LDSL 390

RESULT 14
LRC2 MOUSE
ID LRC2 MOUSE STANDARD; PRT; 371 AA.
AC Q8VDB8: Q9CX04;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing protein 2.
GN LRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Kiss H., Darai E., Kiss C., Kost-Alimova M., Klein G., Dumanaki J.P.,
RA Imreh S.;
RT "Comparative sequence analysis of the common eliminated region 1
RT (C3CER1) on human chromosome 3p21.3 with the orthologous regions on
RT mouse chromosome 9p."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Asburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh S.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC - SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ428068; CAD20990.1; -
CC EMBL; AK010252; BAB26797.1; -
CC MGD; MGI:1921499; Lrrc2.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR_typ.

```

```

DR PFam; PFO0560; LRR; 3.
DR SMART; SM00369; LRR_TYP; 1.
KW Repeat; Leucine-rich repeat.
FT REPEAT 97 120 LRR 1.
FT REPEAT 121 143 LRR 2.
FT REPEAT 144 166 LRR 3.
FT REPEAT 167 189 LRR 4.
FT REPEAT 191 213 LRR 5.
FT REPEAT 214 236 LRR 6.
FT REPEAT 237 259 LRR 7.
FT REPEAT 261 282 LRR 8.
FT REPEAT 283 305 LRR 9.
FT CONFLICT 52 52 Q -> E (IN REF. 2).
SQ SEQUENCE 371 AA; 42955 MW; 55E96FF2D60E6A23 CRC64;

Query Match 9.7%; Score 169; DB 1; Length 371;
Best Local Similarity 25.9%; Pred. No. 7.1e-06;
Matches 70; Conservative 37; Mismatches 91; Indels 72; Gaps 10;

QY 40 VDPVSTLTTPVKTSEFENFKMVTSKKDP-----LS-----KNFPYSLE---HLQTSY- 87
Db 69 VDTSTLLEKIERNSYSRQSRVQKDRKRSFVFEISGTCWKELPDSLSKEQTHLKEWYI 128
QY 88 -CGLVWDVPMCL-KSLRKLDLSHNIKKLPATIGDLHLQELNLDNHNLESFVAL-- 143
Db 129 HSTLIQIIPTYIELFQAMKILDLPKNQITCLPAEIGRLKMLKELNVSNHLKSIPELGD 188
QY 144 -----CHSTLQ-----KSLRSLDLSKNKIKALPVQFCQLQELKMLKLDNE 184
Db 189 CEHLELDLCSGNDLMDLPFELSNLQVTFVDISANKFSVPICVLMRCLQWLDISSN 248
QY 185 LIQFPCKIGQLNLRLSLARKNLPFLPSEFRNLSLEYLDLFGNTFEPQKVPVLKQAP 244
Db 249 LSDLPQDIDRLELQGLFYKKNKLYLPQAMNLK-----K 284
QY 245 LTLLESSARTILHNRIYPYGSHI--IPFHL 272
Db 285 LTLVVS-----GDHLVEVPTALC 303

RESULT 15
LRC2_HUMAN
ID LRC2_HUMAN STANDARD; PRT; 371 AA.
AC Q9BY58; Q961T5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leucine-rich repeat-containing protein 2.
GN LRR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21906202; PubMed=11896456;
RA Kiss H., Yang Y., Kiss C., Andersson K., Klein G., Imreh S.,
RA Dumanaki J.P.;
RT "The transcriptional map of the common eliminated region 1 (C3CER1) in
RT 3p21.3.";
RL Eur. J. Hum. Genet. 10:52-61 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, Kidney, and Stomach;

```



```

QY 241 LQAPLLESSARTILHNRPYSYHII--PFLHLC 272
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 ---KZLLVVS-----GDHLWELTALC 303

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:59:10 / Search time 21 seconds
(without alignments)
1548.226 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745
Sequence: 1 MYVLSPVEFIILQLFIQAI.....PIISYFCSLGCYVNSDMLX 338

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283.5	16.2	428	2 S45361	LRR47 protein - fr
2	218	12.5	717	2 T33295	hypothetical prote
3	208.5	11.9	1495	2 T31434	denin-180 - rat
4	201.5	11.5	1256	2 S60461	gene flightless-I
5	196	11.2	492	2 P86283	hypothetical prote
6	192.5	11.0	886	2 T40734	probable adenylate
7	192.5	11.0	2026	1 OYBY	adenylate cyclase
8	191	10.9	380	2 T01281	probable leucine-r
9	189.5	10.8	526	2 C84532	hypothetical prote
10	189	10.8	1257	2 A88536	protein B0523.5 [i
11	183.5	10.5	277	2 S25770	RSP-1 protein - mo
12	181.5	10.4	549	2 T41744	hypothetical prote
13	181	10.4	540	2 T12704	leucine-rich prote
14	180.5	10.3	277	2 I60122	rsu-1 homolog - hu
15	180.5	10.3	572	2 T30947	hypothetical prote
16	180.5	10.3	613	2 A88684	hypothetical prote
17	179	10.3	559	2 T42998	protein AC7.2 [imp
18	177	10.1	268	2 T45616	Ras-binding protei
19	175	10.0	1839	1 OYBYK	hypothetical prote
20	173	9.9	424	2 S27783	adenylate cyclase
21	173	9.9	902	2 T00588	hypothetical prote
22	172.5	9.9	1119	2 AD1822	hypothetical prote
23	171.5	9.8	404	2 T08549	leucine-rich-repea
24	171	9.8	382	2 T04260	hypothetical prote
25	170	9.7	1039	2 T22117	hypothetical prote
26	166.5	9.5	2493	2 A55481	hypothetical prote
27	166	9.5	268	2 T19697	adenylate cyclase
28	165	9.5	1268	2 A49674	hypothetical prote
29	163.5	9.4	2145	2 JC4747	flightless-1 homol
					adenylate cyclase

30	161.5	9.3	626	2 AE0123	probable antigenic
31	161.5	9.3	853	2 T17461	disease resistance
32	160.5	9.2	904	2 T46170	disease resistance
33	159.5	9.1	431	2 T27904	hypothetical prote
34	159	9.1	227	2 T27905	hypothetical prote
35	156.5	9.0	613	2 T15489	hypothetical prote
36	155	8.9	1029	2 T00712	protein kinase hom
37	154	8.8	1066	2 T15864	hypothetical prote
38	153.5	8.8	307	2 S36779	ribosome-binding p
39	153.5	8.8	371	2 T49908	hypothetical prote
40	153.5	8.8	1068	2 H96769	hypothetical prote
41	153.5	8.8	1232	2 T05322	hypothetical prote
42	153	8.8	1088	2 E86312	hypothetical prote
43	152	8.7	967	2 T48210	FlilA6.9 protein -
44	152	8.7	1011	2 C84524	probable disease r
45	151.5	8.7	720	2 T02361	hypothetical prote

ALIGNMENTS

RESULT 1

S45361
LRR47 protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text_change 2-Jul-2000
C:Accession: S45361; S44046
R:Ntwasa, M.; Buchanan, S.G.S.C.; Gay, N.J.
Biochim. Biophys. Acta 1218, 181-186, 1994
A:Title: Sequence and expression of LRR47, a novel embryonic leucine rich repeat protein
A:Reference number: S45361; MUID:94289476; PMID:8018718
A:Accession: S45361
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-428 <NW>
A:Cross-references: EMBL:X75760; NID:g415946; PIDM:CAA53387.1; PID:g415947
A:Gene: FlyBase:Lrr47
A:Cross-references: FlyBase:FBgn0010398
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match 16.2%; Score 283.5; DB 2; Length 428;
Best Local Similarity 27.7%; Pred. No. 6.6e-14;
Matches 92; Conservative 55; Mismatches 140; Indels 45; Gaps 8;

QY	15	LFQIAISSSLKGFSLAMELAHKG-----CNVDTPVSTLTTPGKTSEFENEKTKWVITSKD	69
DB	31	LLIKCDPIQLKGFLOTLLKGMWDKDAINRLNNAATAIPQAKP----QVRWVTSKRS	146
QY	70	YPLSKNFPYSLEHLQTSYCGLVVDVDMRMCLSKLRKLDLSHNHKKLPATIGDLIHQL	129
DB	147	YPI-KGFPRTLKSLTINNSQLVKLSFEICTLRNLTKLDVSGNKLAKIPSELGEL-PTTSL	204
QY	130	NINNEHLSFS-----VALCHSTLQKLSLSLDSKKIKALPVQFCQLKMLKLD	182
DB	205	HLGNLLGTQNDWCMLRGTKLC-----QSLGELDLSGNGLTFEPPLPVKFSLVSLN	259
QY	183	NELIQCPKIGQLNRLFLSAARNKLPF-PSEFNLSLEYLDLFGNTFE-----QPKV	235
DB	260	NLLSLFATRMKALKKLYVCSNELSESJPAVEDRIDLDLVAGNCKEFNADAQKMC	319
QY	236	LPVTKLQAPLTLESSARTILHNRIPYGSHTIPPHLCQDLDTAKICVGRFCLNS-----	290
DB	320	LQKAASNSPQLWLLGARVDKXKMLPLSAGSIPAVLIDLIREAPRCPCGELCYAQRKEDL	379
QY	291	-----FIQTTTNKLSVAHTVVVDN	312
DB	380	FORVVPQKFSVKNLTSREHQIYADVVLCD	411

RESULT 2

T33295
hypothetical protein F26D11.i1 - *Caenorhabditis elegans*

C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33295
R;Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid F26D11.
A;Reference number: Z21316
A;Accession: T33295
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-717 <GEI>
A;Cross-references: EMBL:AF068716; PIDN: AAC17752.1; GSPDB: GN00023; CESP: F26D11.1
A;Experimental source: strain Bristol N2; clone F26D11
C;Genetics:
A;Gene: CESP:F26D11.1
A;Map position: 5
A;Introns: 34/2; 119/2; 226/1; 302/3; 368/3; 576/2; 623/2

Query Match 12.5%; Score 218; DB 2; Length 717;
Best Local Similarity 39.4%; Pred. No. 1.1e-08;
Matches 61; Conservative 17; Mismatches 69; Indels 8; Gaps 3;

QY 80 LEHLQTSYCGVVRDMRLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNLDNHELSF 139
DB 38 LEDLNLTMMNKKLDHKLFSRLHRLILDVSDNELAVLPABIGNLTQIIELNARN----- 92
QY 140 SVALCHSTLQ--KSLRSLSLSSKIKALPVQFCOLQELKNLKDNDNELIQFPCKIGQLIN 197
DB 93 SIAPKPTQCNCKLLTTLNLSNPPFTLPETICECSITLSLNETSLTLLPSNIGSLTN 152
QY 198 LRFLSAARKLPLPSEFRLNLSLEYLDLFGNTE 231
DB 153 LRVLEARDNLLRTIPLSIVELRKLDELQONELE 187

RESULT 3
T31434
densen-180 - rat
N;Alternate names: postsynaptic density protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C;Accession: T31434
R;Apperson, M.L.; Moon, I.S.; Kennedy, M.B.
J. Neurosci. 16, 6839-6852, 1996
A;Title: Characterization of densen-180, a new brain-specific synaptic protein of the O-
A;Reference number: Z21033; MUID: 96421675; PMID: 8824323
A;Accession: T31434
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1495 <APP>
A;Cross-references: EMBL:U66707; NID: g1657757; PID: g1657758; PIDN: AAC52881.1
A;Experimental source: strain Sprague Dawley
C;Function:
A;Description: supposed to participate in specific adhesion between presynaptic and post

Query Match 11.9%; Score 208.5; DB 2; Length 1495;
Best Local Similarity 34.1%; Pred. No. 1.5e-07;
Matches 58; Conservative 28; Mismatches 75; Indels 9; Gaps 4;

QY 75 NFPSYLSLHQTSCYGLVRVDMRLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNLDN 134
DB 48 NFERTLELYLDANQIEEIPKQFNQALRKISIPDNLSSLTSTASLVNKLDELISKN 107
QY 135 HLESF--SVLCHSTLQSLRSLSLSSKIKALPVQFCOLQELKNLKDNDNELIQFPCKI 192
DB 108 GVQEPENIKCC-----KCLTITEASVNPISKLPGDFTQLNLNLTQLYLNDAFLELPANF 162
QY 193 GOLINLRFLSAARKLPLPSEFRLNLSLEYLDLFGNTEPQKVLPVKIQAPLTLES 240
DB 163 GRVLKRLILELRENHLKTLPKSMHKLQALRELGLGNNEFSELPELVLDIQ 212

RESULT 4

S60461
gene flightless-I protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 2i-Jul-2000
C;Accession: S60461; T08425
R;de Couet, H.G.; Fong, K.S.K.; Weeds, A.G.; McLaughlin, P.J.; Miklos, G.L.G.
Genetics 141, 1049-1059, 1995
A;Title: Molecular and mutational analysis of a gelsolin-family member encoded by the fly
A;Reference number: S60461; MUID: 96129280; PMID: 8582612
A;Accession: S60461
A;Molecule type: DNA
A;Residues: 1-1256 <DEC>
A;Cross-references: GB:AF017777; EMBL:U28044; NID: g3004652; PIDN: AAC28407.1; PID: g3004666
R;Maliszka, R.; de Couet, H.G.; Miklos, G.L.
Proc. Natl. Acad. Sci. U.S.A. 95, 3731-3736, 1998
A;Title: Data transferability from model organisms to human beings: insights from the fur
A;Reference number: Z16415; MUID: 98188272; PMID: 9520435
A;Accession: T08425
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1256 <MAL>
A;Cross-references: GB:AF017777; GB:U28044; NID: g3004652; PIDN: AAC28407.1; PID: g3004666
C;Genetics:
A;Gene: fli-I; flightless-I
A;Cross-references: FlyBase:FBgn0000709
A;Introns: 18/3; 1070/2; 1120/3
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;53-75/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;76-100/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;101-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;124-147/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;148-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;171-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;220-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;243-265/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;266-288/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;497-830/Domain: gelsolin repeat homology <GEL1>
F;892-1250/Domain: gelsolin repeat homology <GEL2>

Query Match 11.5%; Score 201.5; DB 2; Length 1256;
Best Local Similarity 33.7%; Pred. No. 3.9e-07;
Matches 63; Conservative 30; Mismatches 63; Indels 31; Gaps 6;

QY 75 NFPSYLSLHQTSCYGLVRVDMRLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNLDN 134
DB 210 NFPTSIDSLAN-----LC-----ELDLSNLSPLKLPDCVYVNVTLVRLNLSN 252
QY 135 HLESFVALCHSTLQSLRSLSLSSKIKALPVQFCOLQELKNLKDNDNEL--IQFPCKI 192
DB 253 ELTELTAQV---ELWQRLSLSLNRNLQALVALPAALCKLPKRLRLVNDKLNFEIPSGI 309
QY 193 GOLINLRFLSAARKLPLPSEF--RNLSEYLDLFGNTEPQKVLPVKIQAPLTLES 251
DB 310 GKLGALVFSANLLEWPEGCRCGALKQLNLSNR-----LITPDALHLESEGL 361
QY 252 ARTILEN 258
DB 362 DQDLN 368

RESULT 5
F86263
hypoetical protein F3K23.23 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: F86263
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huzar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

```

QJ 971 GDLEMTDLRLNLYNRYRISIKTNASKJONLFLTDNRISNFEDNLPKURALEQENPIT 1030
DB

```


A: Introns: 199/1; 503/3
A: Note: F15JL1.40

Query Match 10.4%; Score 181.5; DB 2; Length 549;
Best Local Similarity 32.8%; Pred. No. 4.5e-06;
Matches 62; Conservative 24; Mismatches 86; Indels 17; Gaps 6;

QY 100 LKSLRKLDLSHNHIKKLPATIGDLIHLQELINLNDNHLSFVALCHSTLQKSLRSLDLSK 159
DB 268 LSSLTKLDLSHNRIGQLPESIGELLNLVYLNLSGNQLSLPSAF--SLAVR-DEEDLSC 324
QY 160 KTKIKALPVQFCQLOELKNLKLDELNIQPPCKTGQLINLRFLSAARKLPFLPSEFRNL- 218
DB 325 NNLPIIPESIGSLVSLKLDVETNDEEIPYSIGGSSLIETLRADYNKALKALPEAIGKIT 384
QY 219 SLEVLDFGNTFQPKVLPVLIKQAPLITLESARTILHNRIYGS-HIIPFHLCDLDT 277
DB 385 LLEILSVRYNIR-----QLPTM--SSLASLKELDVSNELSVFSLCFATTL 432
QY 278 AKICVCGRF 286
DB 433 VKLIGNNF 441

RESULT 13
T12704
leucine-rich protein - common sunflower
C: Species: Helianthus annuus (common sunflower)
C: Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C: Accession: T12704
R: Evrard, J.L.
submitted to the EMBL Data Library, September 1994
A: Reference number: Z17575
A: Accession: T12704
A: Status: preliminary; translated from GS/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-540 <EVR>
A: Cross-references: EMBL:X81997; NID:e977240; PID:e120583
C: Genetics:
A: Gene: SPL7

Query Match 10.4%; Score 181; DB 2; Length 540;
Best Local Similarity 29.0%; Pred. No. 4.8e-06;
Matches 90; Conservative 42; Mismatches 122; Indels 56; Gaps 15;

QY 12 LQLLFTQALSSSLKFLSAMRLAHRCNVDTVSTLTP-----VKTSEFENFKMVIITS 66
DB 166 LQKLYDTAKSA--GKLSAL-----DKRVSSMLPTTAEVVK-EEADILQDALMNG 212
QY 67 KKDYPILS-KNFXY-----SLEHQTSYCGSLVRVDMRLCLKSLRKLDLSHNHIKKL 116
DB 213 VERIDLSRRRLFPVPEAFQKMLTVSLDLSNNKLTAIPESLAGLTSLEELNLSANLFESL 272
QY 117 PATIGDLIHLQELINLNDNHLSFVALCHSTLQKSLRSLDLSKNKIKALPVQF-COLOEL 175
DB 273 PQTIGSLQHLQELNLSVRNKLTLSPDGICK--RSILELDASFNQITLYLPANIGYGLNL 329
QY 176 KNLKLDDELIIQPPCKIGQILNLRFLSAARKLPFLPSEFRNL-SLEVLDFGNTFQPK 234
DB 330 KKLIMPLNVRSLPTSIGEMISQVLDVHFNTLRGPPSIGMKLXLEVLNL-GSNFNDF 388
QY 235 VLP-----VKKL-----QAPLTI--LESSARTL-HNRIPYGSHTIPFHLCDL 275
DB 389 ALPETIGSLTRLEIDICNNQIQQLPITFGRVLSITRLVDEHNPLTVSP---PEVVAE 445
QY 276 DTAKICVCGR 285
DB 446 EAVKVTMSKR 455

RESULT 14
I60122
isu-1 homolog - human

C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 05-Nov-1999
C;Accession: I60122
R;Cutler, M.L.; Bassin, R.H.; Zanoni, L.; Talbot, N.
Mol. Cell. Biol. 12, 3752-3756, 1993
A;Title: Isolation of rsp-1, a novel CDNA capable of suppressing v-Ras transformant.
A;Reference number: I60122
A;Accession: I60122
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-277 <RES>
A;Cross-references: GB:L12535; NID:G434050; PIDN:AAA60292.1; PID:G434051
C;Genetics:
A;Gene: RSU-1
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match 10.3%; Score 180.5; DB 2; Length 277;
Best Local Similarity 30.7%; Pred. NO. 2.2e-06;
Matches 59; Conservative 27; Mismatches 81; Indels 25; Gaps 3;

Qy 72 LSKNPPYSLEH--LQTSYCGLVRYDMRMLCKLSRKLDLSHNHKKLPATIGDLIHLQE 128
Db 31 LDVNGLFSLSHITQVLVSHNKLTMPPNIAELKNLEVLNFFNQIEELPTQISLQKLKH 90
Qy 129 LNLNDNHLESTF-----VALCHSTLQX-----SLRSLDSKNKIKALP 166
Db 91 LNLGNRLNTLPRGFGSLPALEVLDTLVNYSNSLPNGNFFYLTTLRALYLSNDPPEILP 150
Qy 167 VQFQQLKLNKLDNDNELIQFCKIGOLINLRFSSAARNKLPFLPSEPNLSLEYLDLP 226
Db 151 PDIGKLTQLQLSLDRNDLISLPKEIGELTQLKELHQGNRLTWLPPELGNLDLTQOKQV 210
Qy 227 GNTPEQPKVLPV 238
Db 211 FEAENNPWVPI 222

RESULT 15
T30947
hypothetical protein AC7.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30947
R;Neelson, J.; Hawkins, J.
submitted to the EMBL Data Library, June 1999
A;Description: The sequence of A. C. elegans cosmid AC7.
A;Reference number: Z20944
A;Accession: T30947
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-572 <NEL>
A;Cross-references: EMBL:U61957; PIDN:AAB03417.2
A;Experimental source: strain Bristol N2; clone AC7
C;Genetics:
A;Map position: IV
A;Intons: 61/3; 105/3; 161/3; 234/1; 301/3; 375/1; 500/1
A;Note: AC7.2

Query Match 10.3%; Score 180.5; DB 2; Length 572;
Best Local Similarity 24.2%; Pred. NO. 5.6e-06;
Matches 71; Conservative 42; Mismatches 101; Indels 79; Gaps 8;

Qy 11 ILQLLFLQAISSSLK-GFLSAMRLAH-RGCNVDTPVSTLTPVKTSEPFNKMTVTSKK 68
Db 154 LTVIFLQVPVVIKIGSLTILWLRNYIVADEQIGNLSKULQMDVRENKI-----R 206
Qy 69 DYPLSKNPPYSLEHLQTSYCGLVRYDMRMLCKLSRKLDLSHNHKKLPATIGDLIHLQE 128
Db 207 ELPSAIGKLTSLVCLVSYNHLTRVPEIGDCHSLTQLDLQNDLSLSELPVSGIKVNLVR 266
Qy 129 LNLNDN-----HLESTFVALCH-----STLQKLSRLDSKNKIKALPV 167
Db 267 IGIKYNKIRCIPTSELESQQLEEFVESNHQLLPPNLLTLPKIHVTNLSRNLTAFFA 326

```

QY 168 Q-----FCQOELKNUKLDONELIQPPCKIG-----193
DB 327 GPOQOFVSTVINMEHQISOIKIPIGFSKATRLTKLNKENELVSLPLDGMGWSITELN 386
QY 194 -----QNLNLFELGAARNKLPFLPSEFRNL-LEYLDLFGNTE 231
DB 387 LSTNQLKVPLEDIEKVLNLEILVSNQNLKLPQICNLKLNKLELDLENELE 439

Search completed: June 15, 2004, 10:03:20
Job time : 23 secs

```

Search completed: June 15, 2004, 10:03:20
Job time : 23 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 15, 2004, 10:01:25 ; Search time 48 Seconds
(without alignments)
1983.821 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYV-SPVEFIILQLFIQAI.....PIISYFCSLGCYVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1158786 seqs, 281726120 residues

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1649	94.5	416	15 US-10-312-088-44	Sequence 44, Appl
2	1315	75.4	259	15 US-10-312-088-43	Sequence 43, Appl
3	227.5	13.0	444	15 US-10-369-493-20189	Sequence 20189, A
4	218	12.5	717	15 US-10-369-493-6391	Sequence 6391, Ap
5	217	12.4	1489	15 US-10-115-482-18	Sequence 18, Appl
6	214	12.3	231	15 US-10-369-493-11177	Sequence 11177, A
7	201.5	11.5	288	15 US-10-369-493-19830	Sequence 19830, A
8	199.5	11.4	365	12 US-10-087-192-69	Sequence 69, Appl
9	197.5	11.3	546	9 US-09-903-320-250	Sequence 250, App
10	197.5	11.3	546	9 US-09-903-088B-250	Sequence 250, App
11	197.5	11.3	546	9 US-09-905-291A-250	Sequence 250, App
12	197.5	11.3	546	9 US-09-902-853-250	Sequence 250, App
13	197.5	11.3	546	9 US-09-907-824-250	Sequence 250, App
14	197.5	11.3	546	9 US-09-907-841-250	Sequence 250, App
15	197.5	11.3	546	10 US-09-904-011-250	Sequence 250, App

16	197.5	11.3	546	10 US-09-906-742-250	Sequence 250, App
17	197.5	11.3	546	10 US-09-906-838-250	Sequence 250, App
18	197.5	11.3	546	10 US-09-907-613-250	Sequence 250, App
19	197.5	11.3	546	10 US-09-907-942-250	Sequence 250, App
20	197.5	11.3	546	10 US-09-904-859-250	Sequence 250, App
21	197.5	11.3	546	10 US-09-909-204-250	Sequence 250, App
22	197.5	11.3	546	10 US-09-904-820-250	Sequence 250, App
23	197.5	11.3	546	10 US-09-904-786-250	Sequence 250, App
24	197.5	11.3	546	10 US-09-906-646-250	Sequence 250, App
25	197.5	11.3	546	10 US-09-906-700-250	Sequence 250, App
26	197.5	11.3	546	10 US-09-903-786-250	Sequence 250, App
27	197.5	11.3	546	10 US-09-902-903-250	Sequence 250, App
28	197.5	11.3	546	10 US-09-903-749A-250	Sequence 250, App
29	197.5	11.3	546	10 US-09-904-119-250	Sequence 250, App
30	197.5	11.3	546	10 US-09-904-956-250	Sequence 250, App
31	197.5	11.3	546	10 US-09-902-736-250	Sequence 250, App
32	197.5	11.3	546	10 US-09-907-794-250	Sequence 250, App
33	197.5	11.3	546	10 US-09-903-943-250	Sequence 250, App
34	197.5	11.3	546	10 US-09-904-462-250	Sequence 250, App
35	197.5	11.3	546	10 US-09-907-925-250	Sequence 250, App
36	197.5	11.3	546	10 US-09-902-692-250	Sequence 250, App
37	197.5	11.3	546	10 US-09-903-520-250	Sequence 250, App
38	197.5	11.3	546	10 US-09-905-056-250	Sequence 250, App
39	197.5	11.3	546	10 US-09-909-064-250	Sequence 250, App
40	197.5	11.3	546	10 US-09-904-553-250	Sequence 250, App
41	197.5	11.3	546	10 US-09-905-381-250	Sequence 250, App
42	197.5	11.3	546	10 US-09-905-088-250	Sequence 250, App
43	197.5	11.3	546	10 US-09-907-575-250	Sequence 250, App
44	197.5	11.3	546	10 US-09-905-075-250	Sequence 250, App
45	197.5	11.3	546	10 US-09-902-759-250	Sequence 250, App

ALIGNMENTS

```
RESULT 1
US-10-312-088-44
; Sequence 44, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabric, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdock, Paul R.
; APPLICANT: Smith, Randall F.
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Xie, Qing
; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP5C029
; CURRENT APPLICATION NUMBER: US/10/312,088
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-088-44
```

Query Match 94.5%; Score 1649; DB 15; Length 416;
Best Local Similarity 99.1%; Pred No. 8.5e-154;
Matches 320; Conservative 1; Mismatches 0; Indels 2; Gaps 1;


```

Query Match      75.4%; Score 1315; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 3.8e-121;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OQY 1 MYVLSPEVFIIQLLFIQAISSLKGLFSAMLAHRCNCVDTPVSTLTTPVKTSZEFNFKT 60
      |||||
dDb 1 MYVLSPEVFIIQLLFIQAISSLKGLFSAMLAHRCNCVDTPVSTLTTPVKTSZEFNFKT 60

OQY 61 KVVITSKKDYPISKNFPSYLEHLQTSYCGLVRYDVMRLCLKSLRKLDLSNNHKKLPATI 120
      |||||
dDb 61 KVVITSKKDYPISKNFPSYLEHLQTSYCGLVRYDVMRLCLKSLRKLDLSNNHKKLPATI 120

OQY 121 GLLIHLQELNLDNHNLESFVALCHSTLQKLSRLDLSKNKIKALPVQFCQLQELKNLKL 180
      |||||

```

RESULT 4
US-10-369-493-6391
Sequence 6391, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldsman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52352)B

PRIOR APPLICATION NUMBER: 60/296,965

QY Z13 SUEILDU=GNIFEQPKVLI

219 SLEYLDLFGNTFEQPKVLP--VIKLQAPLTL-LESSARTILHNRIPYGSHIIPFHL 271

DB 165 NLETSLYRNQIE---LPPEIGKLENKATLYIDNNKLTILPPPEISELKNLITLNL 217

RESULT 7

US-10-369-493-19830
; Sequence 19830, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19830
; LENGTH: 288
; TYPE: PRT
; ORGANISM: No. US20030233675altoc punctiforme
US-10-369-493-19830

Query Match 11.5%; Score 201.5; DB 15; Length 288;
Best Local Similarity 33.0%; Pred. No. 2.9e-11;
Matches 61; Conservative 29; Mismatches 78; Indels 17; Gaps 6;
QY 64 ITSKKDYPLSKN-----FP---YSLEHLQF---SYGGLVVRVDMRLCLKSLRKLDLSHNI 113
DB 32 LTQLQQLDSNRQVQLPEAIASTLQQLTDLSSNNKLTQLPEAIASTLQQLTDLSSNNQ 91
QY 114 KQIPATIGDLHQLNLDNHNLESFVALCHSTLOKSLRSIDLSKXKIKALPVQFCQ 173
DB 92 TELPEAIASTLQQLNLRNNQTLPEAIASTL---RLQRLDLSNNQTLPEAIASTL 148
QY 174 ELKNKLDNDELIOFCKIGQLINERFLSAARNKLPFLPSEFRLN---SLBYLDLFGNTEPQ 232
DB 149 QLQSFDSLHSHNELTELPNSLRLLYLEIFDCGNSNLRQVPSVIELKGLKELYIYANDLE- 207
QY 233 PKVLP 237
DB 208 --VIP 210

RESULT 8

US-10-087-192-69
; Sequence 69, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(365)

; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-087-192-69

Query Match 11.4%; Score 199.5; DB 12; Length 365;
Best Local Similarity 28.6%; Pred. No. 6.5e-11;
Matches 77; Conservative 29; Mismatches 88; Indels 75; Gaps 8;
QY 72 LSKNPFYSLEHLQTSYCGLVVRVDMRLCLKSLRKLDLSHNIKKLPATIGDLHLQELNL 131
DB 55 LSKN-----RLVEVPMELCCQFVSEILNLXHNCRVIRPEAIVNLQMLCHLNL 101
QY 132 NDHNLESFVALCHSTLOKSLRSIDLSKXKIKALPVQFCQELKNLKLDDNDELIOFCK 191
DB 102 SRNQLSALPACL-----GHLKVLTIASNNKLSLPEEIGQLKQMLDVSCEITALPQ 157
QY 192 IGOJNLNRLSAAARNKLPFLPSEFRLNLSLEYLDLFGNTEFQPKVLPVI---KLQAPUTLL 248
DB 158 IGOJLSRLNLRNVRNRYLVLPPELVDLPLVKDFSCN---KVLVIVPCFREMQLQVLL 214
QY 249 ESSARTILHNRIPIYGHSHIIPFHLQDLOLTAKICVCGRFLNSFIQTTTNNLHSHVAHTV 308
DB 215 ENN-----PLQS-----PPAQICTKGK-----VH----- 233
QY 309 IVDNLGCTEAPILSYFCSLGCYVNSDML 337
DB 234 -----IFKYLSTIQACQIKTSDSL 251

RESULT 9

US-09-909-320-250
; Sequence 250, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/43,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/445,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/446,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08

;; PRIOR APPLICATION NUMBER: PCT/US99/20944
;; PRIOR FILING DATE: 1999-09-13
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 250
;; LENGTH: 546
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-909-320-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;
QY 57 NFKTKWITSKKYPLSKNFFYSLEHQTSGYGLVRVDMRLCLSKSLKLDLS----- 109
Db 312 NEGKGLVLN-----SLKXMANLTELRLRCDLERIPHSFSLNQLQIDKNNLXTI 365
QY 110 -----HMHKKLPATIGDLIHQELNLDNHNLESFSA-CHSTLOKS 151
Db 366 BEIISFQHLRLTCLKWNHAYIPIQIGNLTNLERLYLNENKIEKPTQFYC---RK 422
QY 152 LRSDLSSKNIKALPVQFCQQLQELNKLDDNELLOFPCKIGQLINLRLSARKNKLPL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPELFPCKRLRALHGNVQLSL 482
QY 212 PSEFRNL-SLEYDLFGNTFQPKVLPVYKLAQPLTLLESSARTI---LHNRIIP 261
Db 483 PSRVGELTNLTQIELRGNLE---CLPVELGECF--LXKSLGVVEBDFN-LIP 531

RESULT 10
US-09-909-088B-250
; Sequence 250, Application US/0909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.

;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Mather, Jennie P.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William, I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: 10466-14
;; CURRENT APPLICATION NUMBER: US/09/909,088B
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: US 60/143,048
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: US 60/145,698
;; PRIOR FILING DATE: 1998-07-26
;; PRIOR APPLICATION NUMBER: US 60/146,222
;; PRIOR FILING DATE: 1998-07-28
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1998-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/20944
;; PRIOR FILING DATE: 1999-09-13
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 250
;; LENGTH: 546
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-909-088B-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;
QY 57 NFKTKWITSKKYPLSKNFFYSLEHQTSGYGLVRVDMRLCLSKSLKLDLS----- 109
Db 312 NEGKGLVLN-----SLKXMANLTELRLRCDLERIPHSFSLNQLQIDKNNLXTI 365
QY 110 -----HMHKKLPATIGDLIHQELNLDNHNLESFSA-CHSTLOKS 151
Db 366 BEIISFQHLRLTCLKWNHAYIPIQIGNLTNLERLYLNENKIEKPTQFYC---RK 422
QY 152 LRSDLSSKNIKALPVQFCQQLQELNKLDDNELLOFPCKIGQLINLRLSARKNKLPL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPELFPCKRLRALHGNVQLSL 482
QY 212 PSEFRNL-SLEYDLFGNTFQPKVLPVYKLAQPLTLLESSARTI---LHNRIIP 261

DB 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVVEEDLFTLTP 531

RESULT 11

US-09-905-291A-250
; Sequence 250, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tomas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;
Qy 57 NFKTKVITSKDYPLSKPFYSLEHLQTSYCGVLRVDMELCLKSKRKLDLS----- 109
Db 312 NEGTKLIVLN-----SLKFCANLTELRLCRLERIPHSIFSLHNLQELDKNNLKITI 365
Qy 110 -----HSHKKLPATIGDLIHLQELNLDNHLSEFSVALCHSTLQKS 151
Db 366 EEIISFQHLHRLTCLKLWYXHYAVIPQIGNLNLRLYLNRRNKIKIPQLFYC---RK 422
Qy 152 LRSLDLSKNIKALPVQFCQLQELKNLKLDDNLIOPCKIGQLINRFLSAARKLPLFL 211
Db 423 LRYLDLSHNNLTPLPADIGLLQNLQNLATITANRIETLPPLEFQCKRLRALHGLGNVLQSL 482
Qy 212 PSFRNL-SLEYLDLQNTFEQPKVLPVKLQAPLTLLESSARTI---LHNRIP 261
Db 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVVEEDLFTLTP 531

RESULT 12

US-09-902-853-250
; Sequence 250, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tomas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944

```

; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-250

Query Match      11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

Qy 57 NFKTKMVTSSKDYPLSKNFPYSLEHLOTQSYGVLVRVDMRMCLCKSLRKLKDLSS----- 109
Db 312 NEGTKLIIVN-----SLKMANUTELELRICDLERIPHSIFSLHNLQEIIDLKNNLKTI 365
Qy 110 -----HNHKKLPATIGDLHLQELNLMNDNHLESFVALCHSTLOKS 151
Db 366 BEIISFQHLHRLTCLKWNHYAIPIQIGNLTNLERLYLNRNKIEKPTQLFYC---RK 422
Qy 152 LRSLDLSKXKIKALPVQFCQQLQELKNLKLDDNELTQFCKIGQLINLRFLSAARNKLPFL 211
Db 423 LRYLDLSHNNLTFCPADGLLQNLQNLAIANRIETLPPFLFQCKRLALHGNVQL 482
Qy 212 PSEFNL-SLEYLDLFGNTFQPKVLPVVKLOAPLTLESSARTI---LHNRIP 261
Db 483 PSRVGELTNLTQIELGNRLS---CLPVELGEC2---LLKRSGLVVEEDLFTNLP 531
```

```

RESULT 13
US-09-907-824-250
; Sequence 250, Application US/09907824
; Publication No. US2002019767A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Deanoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
```

```

; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paori, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-250

Query Match      11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

Qy 57 NFKTKMVTSSKDYPLSKNFPYSLEHLOTQSYGVLVRVDMRMCLCKSLRKLKDLSS----- 109
Db 312 NEGTKLIIVN-----SLKMANUTELELRICDLERIPHSIFSLHNLQEIIDLKNNLKTI 365
Qy 110 -----HNHKKLPATIGDLHLQELNLMNDNHLESFVALCHSTLOKS 151
Db 366 BEIISFQHLHRLTCLKWNHYAIPIQIGNLTNLERLYLNRNKIEKPTQLFYC---RK 422
Qy 152 LRSLDLSKXKIKALPVQFCQQLQELKNLKLDDNELTQFCKIGQLINLRFLSAARNKLPFL 211
Db 423 LRYLDLSHNNLTFCPADGLLQNLQNLAIANRIETLPPFLFQCKRLALHGNVQL 482
Qy 212 PSEFNL-SLEYLDLFGNTFQPKVLPVVKLOAPLTLESSARTI---LHNRIP 261
```

Db 483 PSRVGELTNLTQIELGRNLE---CLPVELGECF--LLKRSGLVVEEDLFTLTP 531

RESULT 14

US-09-907-841-250
; Sequence 250, Application US/09907841
; Publication No. US20020198366A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Christopher J.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/907,841

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; Remaining Prior Application data removed - See File Wrapper or PALM.

; SEQ ID NO 250

; LENGTH: 546

; TYPE: PRT

; ORGANISM: Homo Sapien

US-09-907-841-250

Query Match

Best Local Similarity 11.3%; Score 197.5; DB 9; Length 546;

Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTQWVTSKXDPYLSKXNFPYSLEHQTSCYGLVVDMMCLXSLRKLKLS-----109

Db 312 NEGTKLIVN-----SLKKMANTELEIRCDLERIPHSIFSLHNLQELDKDNKTKTI 365
QY 110 -----HHHIKKLPATIGLIHLQELNLDNHELESFVALCHSTLQKS 151
Db 366 ERIISFOHLRLTCLKWYHNIAYIPQIGNLTWLERLYLNKNKIEKPTQLFYC---RK 422
QY 152 LRSIDLKNNKIKALPVQFCQIQELKNLKDNDNELIQFPCKTGQILNRFLSAARNKLPPL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFQCRKURALHGNVLOSL 482
QY 212 PSEFRNL-SLEYLDLFGNTPEQPKVLPVILKQAPLTLESARTI---LNNRIP 261
Db 483 PSRVGELTNLTQIELGRNLE---CLPVELGECF--LLKRSGLVVEEDLFTLTP 531

RESULT 15

US-09-904-011-250

; Sequence 250, Application US/09904011

; Publication No. US20030003530A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,011

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

```
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-250

Query Match      11.3%; Score 197.5; DB 10; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NPKTMMVITSKXDYPLSKNPPYSLEHLQTSYCGLVNDRMMLCKSLRKLKDL----- 109
Db 312 NEGTXLIVLN-----SLKKMANLTELIRCDLERIPHISIFSLHNLQETIDLKNNLKITI 365

QY 110 -----HNHIKKLPATIGDLIHLQELNLDNDNHLESFVALCHSTLQXS 151
Db 366 EBIIISFOHLRLTCLKLWYNHAYIPIQIGNLTNLERLYNKNKIEKIPTQLFYC---RK 422

QY 152 LRSLSKXNKIKALPVQFCQLQELKXNLKDDNELIQPCIKGQLINLRFISAARNKLPFL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRAHLGNVQLQSL 482

QY 212 PSEFENL-SLEYLDLFGNTFQPKVLPVIKQAPLTLLESSARTI---LHNRIIP 261
Db 483 PSRVGELTNLTQIELGRNRL---CLPVELGCP--LLKRSGLVVRBEDLFTIP 531
```

Search completed: June 15, 2004, 10:04:56
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 10:00:10 ; Search time 23 Seconds
(without alignments)
758.677 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MVLSPVEFIILQLFIQAI.....PIISYFSLGLCVNNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCITUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	197.5	11.3	546	4	US-09-907-794A-250
2	197.5	11.3	546	4	US-09-905-125A-250
3	197.5	11.3	546	4	US-09-902-775A-250
4	187.5	10.7	501	4	US-09-907-794A-185
5	187.5	10.7	501	4	US-09-905-125A-185
6	187.5	10.7	501	4	US-09-902-775A-185
7	179	10.3	559	4	US-09-081-149-2
8	178.5	10.2	582	4	US-09-081-149-8
9	177.5	10.2	582	4	US-09-081-149-7
10	162	9.3	1874	4	US-09-331-403-2
11	154	8.8	723	4	US-09-434-408-2
12	152	8.7	292	4	US-09-312-283C-384
13	148.5	8.5	560	3	US-08-592-500-2
14	148.5	8.5	560	3	US-08-195-006-2
15	148.5	8.5	560	3	US-09-063-950-4
16	148.5	8.5	560	5	PCT-US94-07644A-2
17	148.5	8.5	1240	3	US-08-930-996A-4
18	147	8.4	799	3	US-09-180-439-6
19	147	8.4	1112	3	US-09-353-585-2
20	147	8.4	1112	3	US-09-353-585-3
21	142.5	8.2	947	4	US-09-228-986-73
22	141.5	8.1	863	2	US-08-666-271-2
23	140.5	8.1	673	3	US-09-063-950-2
24	140.5	8.1	910	4	US-09-228-986-72
25	138.5	7.9	806	3	US-08-945-983-2
26	138.5	7.9	907	4	US-09-170-496D-264
27	138.5	7.9	907	4	US-09-170-496D-278

ALIGNMENTS

RESULT 1

US-09-907-794A-250

; Sequence 250, Application US/09907794A

; Patent No. 6635468

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Askenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/907,794A

; PRIOR FILING DATE: 2001-07-17

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

Sequence 5, Appli
Sequence 4266, Ap
Sequence 3, Appli
Sequence 3, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 294, App
Sequence 294, App
Sequence 294, App
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 290, App
Sequence 290, App
Sequence 290, App

28 138 7.9 999 2 US-08-473-553A-5
29 137 7.9 115 4 US-09-621-976-4266
30 136.5 7.8 1015 3 US-09-180-439-8
31 134.5 7.7 968 3 US-09-180-439-3
32 134.5 7.7 968 3 US-09-180-439-4
33 133 7.6 1012 2 US-08-475-891A-4
34 133 7.6 1025 2 US-08-567-375-4
35 133 7.6 1025 2 US-08-587-680A-4
36 132 7.6 1119 4 US-09-907-794A-294
37 132 7.6 1119 4 US-09-905-125A-294
38 132 7.6 1119 4 US-09-902-775A-294
39 131 7.5 320 1 US-07-613-083B-1
40 130.5 7.5 1525 3 US-09-191-647-2
41 130.5 7.5 1525 3 US-09-540-245A-2
42 130.5 7.5 1525 3 US-09-540-153-2
43 129.5 7.4 1059 4 US-09-907-794A-290-
44 129.5 7.4 1059 4 US-09-905-125A-290
45 129.5 7.4 1059 4 US-09-902-775A-290

PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 250
LENGTH: 546
TYPE: PRT
ORGANISM: Homo Sapien
US-09-907-794A-250

Query Match
Best Local Similarity 11.3%; Score 197.5; DB 4; Length 546;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NEKTMVTSKDDYSLKSNFYSLEHLQTSYGLVAVDMRMLCLSKRLKDLG----- 109
DB 312 NEGKLVILN-----SLKMANLTELIRCDLERIPHSIFSLHNLQELDKNNLKTII 365
QY 110 -----HNHKKLPATIGDLHLQELNNDHLSFVALCHSTLQKS 151
DB 366 EIIISFQHLHRLTCKLWYNHAYIPQIGNUTNLRGLYNNRKNKIEKIPQFYC---RK 422
QY 152 LRSLSLKNKIKALPVQFCOLQELKNLKDDELIOFPCKIGQLINRLFSLAARNKLPL 211
DB 423 LRYLDLSHNNLTFPADIGLLQNLQNLAITANRIETLPPELFCQCKLRALHGLNVLQSL 482
QY 212 PSEFRNL-SLEVLDLFGNTFEPQKVLVTKLQAPLTLESSARTI---LHNRIIP 261
DB 483 PSRVGELTNLTQIELGRNLE---CLPVELGRCP--LAKRSGLWVEDLFTNIP 531

RESULT 2

US-09-905-125A-250
Sequence 250, Application US/09905125A
Patent No. 6664376
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,125A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 250
LENGTH: 546
TYPE: PRT
ORGANISM: Homo Sapien
US-09-905-125A-250

Query Match
Best Local Similarity 11.3%; Score 197.5; DB 4; Length 546;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NEKTMVTSKDDYSLKSNFYSLEHLQTSYGLVAVDMRMLCLSKRLKDLG----- 109
DB 312 NEGKLVILN-----SLKMANLTELIRCDLERIPHSIFSLHNLQELDKNNLKTII 365
QY 110 -----HNHKKLPATIGDLHLQELNNDHLSFVALCHSTLQKS 151
DB 366 EIIISFQHLHRLTCKLWYNHAYIPQIGNUTNLRGLYNNRKNKIEKIPQFYC---RK 422
QY 152 LRSLSLKNKIKALPVQFCOLQELKNLKDDELIOFPCKIGQLINRLFSLAARNKLPL 211
DB 423 LRYLDLSHNNLTFPADIGLLQNLQNLAITANRIETLPPELFCQCKLRALHGLNVLQSL 482
QY 212 PSEFRNL-SLEVLDLFGNTFEPQKVLVTKLQAPLTLESSARTI---LHNRIIP 261
DB 483 PSRVGELTNLTQIELGRNLE---CLPVELGRCP--LAKRSGLWVEDLFTNIP 531

RESULT 3

US-09-902-775A-250

```

Query Match          11.3%; Score 197.5; DB 4; Length 546;
Best Local Similarity 29.1%; Pred. No. 3.5e-12;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY  57  NFKTKWVITSKKDYPLSKNFPYPYSLEHLQTSYGLVQVDMRLCLKSLRKLDLS----- 109
Db    312  NECTKLIVLN-----SLKKMANLTELRLRCDELRIPHSIFSLHNQLQELDKNNLKTI 385

QY  110  -----NNHIKKLPATIGDLHLQLQELNLDNHHLSFSVALCHSTLQKS 151
Db    366  EETISFQHLRLTCLKLVNHYAVIDIQNLNTLRLYLNRRNKIEKIPQLFVC---RK 422

QY  152  LRSLLDSKKIKALPVQFQQLQELNKLKDDNELIOFPCKIGQILNIRFSLAARNKLPFL 211
Db    423  LRYLDLSHNNLTFLPADIGLLQNLQNLATANRIETLPPELFOCKLRALHLCNNVLQSL 482

QY  212  PSEFRNL-SLEVDLDFGNTFEQPKVLPVTKLOAPLTLLSSARTI---LHNRIIP 261
Db    483  PSRVGLTNTLQELRGNRL-----CLPVELGECF--LLKRSGLVVEEDIENTLP 531

RESULT 4
US-09-907-794A-185
; Sequence 185, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, LLC
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoletti, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Acids and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15

```

APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,125A
CURRENT FILING DATE: 2001-07-12
PCT FILING DATE: 2000-02-22
PCT FILING DATE: 1999-07-07
PCT FILING DATE: 1999-07-26
PCT FILING DATE: 1999-07-28
PCT FILING DATE: 1999-09-08
PCT FILING DATE: 1999-09-13
PCT FILING DATE: 1999-09-15
PCT FILING DATE: 1999-09-15
PCT FILING DATE: 1999-09-15
PCT FILING DATE: 1999-10-05
PCT FILING DATE: 1999-11-29
PCT FILING DATE: 1999-11-30
PCT FILING DATE: 1999-12-02
PCT FILING DATE: 1999-12-02
PCT FILING DATE: 1999-12-02
PCT FILING DATE: 1999-12-16
PCT FILING DATE: 1999-12-20
PCT FILING DATE: 1999-12-20
PCT FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 185
LENGTH: 501
TYPE: PRT
ORGANISM: Homo sapiens
US-09-907-794A-185

Query Match 10.7%; Score 187.5; DB 4; Length 501;
Best Local Similarity 27.2%; Pred. No. 3.6e-11;
Matches 64; Conservative 45; Mismatches 87; Indels 39; Gaps 7;
QY 28 LSAMRLAHGCVNDTPVSTLTTPVKTSS---EFENFKTKVITSKDYPLSKNFPYSLEHL 83
DB 230 LRHLKILHVKSNTKVPNSITDVAPHLTKLVHNDGTLLVLS---LKKQM--NVAEL 283
QY 84 QTSYGLVRVDMRMLCKLSRLKDLSS---EHNHKKLPA 118
DB 284 ELQNCLELERIPHAIFSLNQLQELDLKSNIRITIEIISFQHLKRLTCLKLNKIVTIPP 343
QY 119 TIGDLIHLQELNLDNHLSEFVALCHSTLOKSLRSLDSKNKIKALPVQFQOLQELKML 178
DB 344 SITHVKNLESYFSSNNKLESFPVAV--FSLQK-LRCLDVSNNISMIPIEIGLLQNLQHL 400
QY 179 KLDNDELQFPCKIGQLINLRLPFLSARNKLPPLPSEFRNLS-LEYLDLFGNTFEQ 232
DB 401 HITGNKVDILPKQFKICKRLNLNGNCITSIPKVGQSLQTLQELKGNCLDR 455

RESULT 5
US-09-905-125A-185
Sequence 185, Application US/09905125A
Patent No. 6664376
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin J.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

Query Match 10.7%; Score 187.5; DB 4; Length 501;
Best Local Similarity 27.2%; Pred. No. 3.6e-11;
Matches 64; Conservative 45; Mismatches 87; Indels 39; Gaps 7;
QY 28 LSAMRLAHGCVNDTPVSTLTTPVKTSS---EFENFKTKVITSKDYPLSKNFPYSLEHL 83
DB 230 LRHLKILHVKSNTKVPNSITDVAPHLTKLVHNDGTLLVLS---LKKQM--NVAEL 283
QY 84 QTSYGLVRVDMRMLCKLSRLKDLSS---EHNHKKLPA 118
DB 284 ELQNCLELERIPHAIFSLNQLQELDLKSNIRITIEIISFQHLKRLTCLKLNKIVTIPP 343
QY 119 TIGDLIHLQELNLDNHLSEFVALCHSTLOKSLRSLDSKNKIKALPVQFQOLQELKML 178
DB 344 SITHVKNLESYFSSNNKLESFPVAV--FSLQK-LRCLDVSNNISMIPIEIGLLQNLQHL 400
QY 179 KLDNDELQFPCKIGQLINLRLPFLSARNKLPPLPSEFRNLS-LEYLDLFGNTFEQ 232
DB 401 HITGNKVDILPKQFKICKRLNLNGNCITSIPKVGQSLQTLQELKGNCLDR 455

RESULT 6
US-09-902-775A-185
Sequence 185, Application US/09902775A


```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-081-149-8

Query Match
Best Local Similarity 10.2%; Score 178.5; DB 4; Length 582;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLKLDLSHHHKKLPATIGDLHLQELNLDNHLSEFSVALCHSTOKLSRSLDLSKNK 161
Db 403 SMVELNLATNQLTKIEDVSGLVSLVLEVLISNNLLKKLPGL--GNLRK-LRELDLEENK 459

QY 162 IKALPVQFQLOELKRLKLDNDELIOFPCKIGOLINRFLSAARNKLPFLPSEFRNL-SL 220
Db 460 LESLPNEIAYLKDQLKLVLTNNQLTTLPRGIGHLTNLTJHLGLGENULLTHLPEEIGTLENL 519

QY 221 EYLDLFGNTFEQPKVLPVLIKQAPLTLLSSARTILNHRIPYGSHIIPFH--LCQDLDTA 278
Db 520 EELYLNDN-----PNL-----HSLPPELALCSKLSIM 546

QY 279 KICVC 283
Db 547 SIENC 551

RESULT 9
US-09-081-149-7
; Sequence 7, Application US/09081149A
; Patent No. 6506889
; GENERAL INFORMATION:
; APPLICANT: Han, Min
; APPLICANT: Sieburth, Derek
; TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
; FILE REFERENCE: UTC-02938
; CURRENT APPLICATION NUMBER: US/09/081,149A
; CURRENT FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURES:
; OTHER INFORMATION: Xaa at position 10 can be any amino acid.
US-09-081-149-7

Query Match
Best Local Similarity 10.2%; Score 177.5; DB 4; Length 582;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLKLDLSHHHKKLPATIGDLHLQELNLDNHLSEFSVALCHSTOKLSRSLDLSKNK 161
Db 403 SMVELNLATNQLTKIEDVSGLVSLVLEVLISNNLLKKLPGL--GNLRK-LRELDLEENK 459

QY 162 IKALPVQFQLOELKRLKLDNDELIOFPCKIGOLINRFLSAARNKLPFLPSEFRNL-SL 220
Db 460 LESLPNEIAYLKDQLKLVLTNNQLTTLPRGIGHLTNLTJHLGLGENULLTHLPEEIGTLENL 519

QY 221 EYLDLFGNTFEQPKVLPVLIKQAPLTLLSSARTILNHRIPYGSHIIPFH--LCQDLDTA 278
Db 520 EELYLNDN-----PNL-----HSLPPELALCSKLSIM 546

QY 279 KICVC 283
Db 547 SIENC 551

RESULT 10
US-09-331-403-2
; Sequence 2, Application US/09331403
```

```
; Patent No. 6489147
; GENERAL INFORMATION:
; APPLICANT: ALTMANN-JOHL, Regula; PHILIPPSEN, Peter; ALTHOFER,
; Henning; SEUBERGER, Harald.
; TITLE OF INVENTION: Adenylate cyclase gene, and its use
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM AT-compatible, Pentium III processor
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/331.403
; FILING DATE: 21-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/07309
; FILING DATE: 29-DEC-1997
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-331-403-2

Query Match
Best Local Similarity 9.3%; Score 162; DB 4; Length 1874;
Matches 60; Conservative 31; Mismatches 85; Indels 38; Gaps 7;

QY 70 YPLSNFPYSLEHLQTSYCGLVFVDMR-----MLCLKSLRKLDLSHHHKKLPATI 120
Db 762 YNKIRSLPDSMNQLOK---LAKINLSNNRITHVNDLSKMTSLRTLDLRNRIESIKRV 817

QY 121 GDLIHLQELNLDNHLSEF-----SVALCHSTLQKSLSLDLSKNKIK 163
Db 818 P---NLQNLFLTENRTMFDQDLMLRTLELQNPISLTLKNDYDEHLSTLSISKAKLA 874

QY 164 ALPVQFC-QLOELKRLKLDNDELIOFPCKIGOLINRFLSAARNKLPFLPSEFRNL-SLE 221
Db 875 VLPESLRLRLPRLEKLELSENSITVLPDIKHLKLVHLSVAKNKLSELPDEIAS-KNLX 934

QY 222 YLDLFGNTFEQPKVLPVLIKQAPLTLLSSARTI 255
Db 935 MLDLHCNNL---MTLPAALSTLSLTFVNISSNNML 965

RESULT 11
US-09-434-408-2
; Sequence 2, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
```

```

;
;
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
;

```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-195-006-2

Query Match 8.5%; Score 148.5; DB 3; Length 560;
Best Local Similarity 29.6%; Pred. No. 5.6e-07;
Matches 71; Conservative 30; Mismatches 82; Indels 57; Gaps 12;
QY 29 SAMRLAHRCNVDPVSTLTPTVKTSEFENFKTMVITSKKDYPISKNFYPSLEHLQTSYC 88
DB 71 SGMVQLRMISDSHISAVAPGTFSDLIKLT-----LRLSRN---KITHLPGL- 117
QY 89 GLVRVDMRMCLKSLRKLDLSHNHKKLPATI-GDLIHQLQELNLDNHNHLESFVSVALCHST 147
DB 118 --LQKWL---LQQLFDHNAIRGIDQNMFKLVNQLQELALNQQLDFLPASL--FT 167
QY 148 LQKSLRSIDLKSKIKALPVQFCOLQ-ELKNLKLDDNELI-----QF-- 188
DB 168 NLENLKLDSLGNLTHLPKGLGAQAQKLERLLHSNRLVSLDSGLNSLGLTALTELQFHR 227
QY 189 -----PCKIGQLINLRFLSAARNKLPPLPSEF---RNLSLEYLDLFGNTF-EQPKVL 236
DB 228 NHIRSIAPGAFDRPLNLSLTLSENHLAFPSALFLSHNLT--LTLFENPLAELPGVL 285

RESULT 15
US-09-063-950-4
Sequence 4, Application US/09063950C
Patent No. 6225085
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE OF INVENTION: THEREFOR
FILE REFERENCE: MEI-019
CURRENT APPLICATION NUMBER: US/09/063,950C
CURRENT FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 560
TYPE: PRT
ORGANISM: Homo sapiens
US-09-063-950-4

Query Match 8.5%; Score 148.5; DB 3; Length 560;
Best Local Similarity 29.6%; Pred. No. 5.6e-07;
Matches 71; Conservative 30; Mismatches 82; Indels 57; Gaps 12;
QY 29 SAMRLAHRCNVDPVSTLTPTVKTSEFENFKTMVITSKKDYPISKNFYPSLEHLQTSYC 88
DB 71 SGMVQLRMISDSHISAVAPGTFSDLIKLT-----LRLSRN---KITHLPGL- 117
QY 89 GLVRVDMRMCLKSLRKLDLSHNHKKLPATI-GDLIHQLQELNLDNHNHLESFVSVALCHST 147

Db 118 ----LQKWL---LQQLFDHNAIRGIDQNMFKLVNQLQELALNQQLDFLPASL--FT 167
QY 148 LQKSLRSIDLKSKIKALPVQFCOLQ-ELKNLKLDDNELI-----QF-- 188
DB 168 NLENLKLDSLGNLTHLPKGLGAQAQKLERLLHSNRLVSLDSGLNSLGLTALTELQFHR 227
QY 189 -----PCKIGQLINLRFLSAARNKLPPLPSEF---RNLSLEYLDLFGNTF-EQPKVL 236
DB 228 NHIRSIAPGAFDRPLNLSLTLSENHLAFPSALFLSHNLT--LTLFENPLAELPGVL 285

Search completed: June 15, 2004, 10:03:55
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:54:39 ; Search time 60 Seconds
(without alignments)
1591.684 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYVLSPEVFIIQLLFQAI.....PIISYFCSIGCVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1745	100.0	338	4	AAB27231 Human EXM
2	1660	95.1	414	4	AAM40049 Human pol
3	1649	94.5	416	5	AAE17323 Human leu
4	1632	93.5	429	4	AAM41835 Human pol
5	1606	92.0	308	5	ABG31598 Fruit fly
6	1315	75.4	259	5	AAE17322 Human leu
7	1137	65.2	259	4	AAU20511 Human sec
8	284.5	16.3	428	4	ABE62072 Drosophila
9	218.5	12.5	389	6	ABO14780 Novel hum
10	217	12.4	1489	6	ABO14779 Novel hum
11	217	12.4	1489	6	ABU11038 Human pro
12	211.5	12.1	431	4	ABB61551 Drosophila
13	207.5	11.9	378	4	ABB65454 Drosophila
14	207	11.9	849	4	ABB64971 Drosophila
15	201.5	11.5	1060	4	ABB67403 Drosophila
16	201.5	11.5	1066	4	ABB58166 Drosophila
17	200	11.5	1052	3	ABAB42389 Human ORF
18	198	11.3	761	6	ABO00758 Polypepti
19	198	11.3	761	7	ABE09428 Novel pro
20	197.5	11.3	474	4	AAE19199 Human pro
21	197.5	11.3	546	2	AAE13386 Amiro aci
22	197.5	11.3	546	3	ACJ78562 Human PRO
23	197.5	11.3	546	4	AAE80254 Human PRO
24	197.5	11.3	546	4	AAU29229 Human PRO
25	197.5	11.3	546	5	AAU83653 Human PRO

26	197.5	11.3	546	5	ABE84836 Human PRO
27	197.5	11.3	546	5	ABE95442 Human ang
28	197.5	11.3	546	6	ABU58605 Human PRO
29	197.5	11.3	546	6	ABU71632 Human PRO
30	197.5	11.3	546	6	ABU88153 Novel hum
31	197.5	11.3	546	6	ABU84468 Human sec
32	197.5	11.3	546	6	ABR66342 Human sec
33	197.5	11.3	546	6	ABR65732 Human sec
34	197.5	11.3	546	6	ABU99672 Human sec
35	197.5	11.3	546	6	ABU82911 Human PRO
36	197.5	11.3	546	6	ABU90032 Novel hum
37	197.5	11.3	546	6	ABU71487 Human PRO
38	197.5	11.3	546	6	ABR68281 Human sec
39	197.5	11.3	546	6	ABU96334 Novel hum
40	197.5	11.3	546	6	ABU92765 Human sec
41	197.5	11.3	546	6	ABU80800 Human PRO
42	197.5	11.3	546	6	ABO08842 Human sec
43	197.5	11.3	546	6	ABO02894 Human sec
44	197.5	11.3	546	6	ABR75048 Human sec
45	197.5	11.3	546	6	ABR94810 Human sec

ALIGNMENTS

RESULT 1

AAB27231
ID AAB27231 standard; protein; 338 AA.

XX AAB27231;

XX 27-MAR-2001 (first entry)

XX Human EXMAD-9 SEQ ID NO: 9.

XX Extracellular matrix and adhesion-associated protein; EXMAD; cancer;
inflammation; reproductive disorder; cardiovascular disorder;
immune disorder; musculoskeletal disorder; developmental disorder;
gastrointestinal disorder; cell proliferation disorder.

OS Homo sapiens.

PN WO200606380-A2.

XX 16-NOV-2000.

XX 10-MAY-2000; 2000WO-US012811.

XX 11-MAY-1999; 99US-0133643P.

XX 23-AUG-1999; 99US-0150409P.

XX (INCY-) INCYTE GENOMICS INC.

XX Bandman O, Hillman JL, Tang YT, Lal P, Yue H, Baughn MR, Lu DAM;
Azimzai Y;

XX WPI; 2001-007395/01.

XX N-PSDB; AAC66898.

XX Isolated polynucleotide encoding extracellular matrix or adhesion-associated protein (EXMAD) useful for diagnosing, treating, or preventing disorders associated with expression of EXMAD such as proliferative, immune and genetic disorders.

XX Claim 1; Page 96-97; 129pp; English.

XX The present invention provides the protein and coding sequences for 25 novel extracellular matrix and adhesion-associated proteins (EXMADs). These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5, EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12, EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19, EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are useful in the prevention and treatment of cancers, cell proliferation,

CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA157798-AA161369) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 429 AA;

Query Match 93.5%; Score 1632; DB 4; Length 429;
 Best Local Similarity 99.7%; Pred. No. 2.7e-157;
 Matches 315; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 18 QATSSSLKGLSAMRLAHRGCVDTPTVSTLTPTVKTSSEFENFKTKWITSKDYPLSKNFP 77
 Db 111 KANSSSLKGLSAMRLAHRGCVDTPTVSTLTPTVKTSSEFENFKTKWITSKDYPLSKNFP 170
 QY 78 YSLEHLQTSYCGLVVRVDMRMCLSKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHLE 137
 Db 171 YSLEHLQTSYCGLVVRVDMRMCLSKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHLE 230
 QY 138 SFSVALCHSTLQKSLSLDKSNKKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLIN 197
 Db 231 SFSVALCHSTLQKSLSLDKSNKKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLIN 290
 QY 198 LRFLSAARNKLPPLPSEFRNLSEYLDLFGNTFQPKVLPVKLOAPLTLLSSARTILH 257
 Db 291 LRFLSAARNKLPPLPSEFRNLSEYLDLFGNTFQPKVLPVKLOAPLTLLSSARTILH 350
 QY 258 NRIPYSGSHIIPFLCQDLDTAKICVGRFCNSFIQGTITMNLHSAVHTVVLVDNLGGTE 317
 Db 351 NRIPYSGSHIIPFLCQDLDTAKICVGRFCNSFIQGTITMNLHSAVHTVVLVDNLGGTE 410
 QY 318 APIISYFCSLGCVNNSDM 336
 Db 411 APIISYFCSLGCVNNSDI 429

RESULT 5
 ABG31598
 ID ABG31598 standard; protein; 308 AA.
 AC ABG31598;

DT 05-NOV-2002 (first entry)

DE Fruit fly LRR47 polypeptide 47-33.88.

KW Fruit fly; LRR47 polypeptide 47-33.88; embryonic development deformity;
 KW tumour; diabetes; menstrual disorder; peptide ulcer; arrhythmia; anaemia;
 KW epilepsy.

OS Drosophila sp.

PN CN1341640-A.

PD 27-MAR-2002.

PF 05-SEP-2000; 2000CN-00125025.

PR 05-SEP-2000; 2000CN-00125025.

PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.

PI Mao Y, Xie Y;
 XX WPI; 2002-520716/56.
 DR N-PSDB; ABK90911.
 XX A fruit fly LRR47 polypeptide 47-33.88, useful for curing e.g. tumors and
 PT diabetes.
 XX Claim 1; Page 26-27 (Disclosure); 3pp; Chinese.
 XX The present invention relates to a new fruit fly LRR47 polypeptide 47-
 CC 33.88. The polypeptide is useful for curing several diseases, such as
 CC embryonic development deformity, tumour, diabetes, menstrual disorder,
 CC peptide ulcer, arrhythmia, anaemia and epilepsy. The present amino acid
 CC sequence represents the fruit fly LRR47 polypeptide 47-33.88 of the
 CC invention
 XX
 SQ Sequence 308 AA;

Query Match 92.0%; Score 1606; DB 5; Length 308;
 Best Local Similarity 130.0%; Pred. No. 7.5e-155;
 Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 MRLAHRGCVDTPTVSTLTPTVKTSSEFENFKTKWITSKDYPLSKNFPYSLEHLQTSYCG 90
 Db 1 MRLAHRGCVDTPTVSTLTPTVKTSSEFENFKTKWITSKDYPLSKNFPYSLEHLQTSYCG 60
 QY 91 VRVDMRMCLSKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHLESFVALCHSTLQK 150
 Db 61 VRVDMRMCLSKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHLESFVALCHSTLQK 120
 QY 151 SLRSLDKSNKKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLINLRFSAARNKLP 210
 Db 121 SLRSLDKSNKKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLINLRFSAARNKLP 180
 QY 211 LPSEFRNLSEYLDLFGNTFQPKVLPVKLOAPLTLLSSARTILHNRIPYSGSHIIPFH 270
 Db 181 LPSEFRNLSEYLDLFGNTFQPKVLPVKLOAPLTLLSSARTILHNRIPYSGSHIIPFH 240
 QY 271 LCQDLDTAKICVGRFCNSFIQGTITMNLHSAVHTVVLVDNLGGTEAPIISYFCSLGCV 330
 Db 241 LCQDLDTAKICVGRFCNSFIQGTITMNLHSAVHTVVLVDNLGGTEAPIISYFCSLGCV 300
 QY 331 VNSSDMLK 338
 Db 301 VNSSDMLK 308

RESULT 6
 AAE17322
 ID AAE17322 standard; protein; 259 AA.
 AC AAE17322;

DT 19-APR-2002 (first entry)

DE Human leucine rich repeat protein, sbg442358PROA #1.

KW Human; therapy; wound healing disorder; vaccine; cancer; infection;
 KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
 KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
 KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
 KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
 KW depression; cardiovascular disease; myocardial infarction; renal failure;
 KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
 KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
 KW hyperlipidemia; renal disease; hypoglycaemia; gastrointestinal disease;
 KW neoplasia; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
 KW haemostatic; vulvar; anticonvulsant; antirheumatic; neuroprotective;
 KW nephrotropic; hypotensive; vasotropic; cytoskeletal; cerebroprotective;
 KW allergy; leucine rich repeat protein.

OS Homo sapiens.

XX WO200198342-A1.
 XX 27-DEC-2001.
 XX 22-JUN-2001; 2001WO-US019929.
 XX 22-JUN-2001; 2000US-0213156P.
 XX 22-JUN-2001; 2000US-0213161P.
 XX (SMK) SMITHKLINE BEECHAM CORP.
 XX (SMK) SMITHKLINE BEECHAM PLC.
 XX (GLAX) GLAXO GROUP LTD.
 XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;
 XX Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;
 XX WPI: 2002-139783/18.
 XX N-PSDB; AAD27817.
 XX Novel secreted and membrane-associated polypeptides and polynucleotides
 XX useful for preventing, ameliorating or correcting dysfunction or disease
 XX including diabetes, cancer, hypertension and growth abnormalities.
 XX Claim 1: Page 134; 138pp; English.
 XX The invention relates to secreted and membrane-associated polypeptides
 XX and polynucleotides. The sequences of the invention are useful in
 XX diagnostic assays for detecting diseases associated with inappropriate
 XX activity or levels of these polynucleotides, and in identifying their
 XX agonists and antagonists that are potentially useful in therapy. The
 XX sequences of the invention are useful as vaccines for inducing
 XX immunological responses. The sequences of the invention are useful for
 XX treating cancers, infections, autoimmune disorders, haematopoietic
 XX disorders, wound healing disorders, cholesterol ester storage disease,
 XX inflammation, congenital muscular dystrophy, junctional epidermolysis
 XX bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,
 XX viral and bacterial infections, Alzheimer's disease, asthma, arthritis,
 XX allergies, schizophrenia, sbg44245PR0a-associated disorders,
 XX septicemia, psoriasis, inflammatory bowel disease, transplant rejection,
 XX graft versus host disease, ischaemia, stroke, acute respiratory disease,
 XX syndrome, restenosis, ischaemia, stroke, acute respiratory disease,
 XX brain disorders including paraneoplastic palsy, myotonic dystrophy,
 XX depression, anxiety disorders and sleep disorders, cardiovascular
 XX diseases including congestive heart failure and myocardial infarction,
 XX respiratory diseases including chronic obstructive pulmonary disease,
 XX acute bronchitis and adult respiratory distress syndrome, liver disorders
 XX including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral
 XX and non-viral hepatitis, type II diabetes mellitus, renal disease
 XX including acute and chronic renal failure, glomerulonephritis, Fanconi's
 XX syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia
 XX and tendinitis, gastrointestinal diseases including intestinal
 XX obstruction and tropical sprue, spleen disorders including hypersplenism,
 XX Hodgkin's disease and malignant lymphoma, testicular cancer, male
 XX reproductive diseases including low testosterone and male infertility.
 XX The present sequence is human leucine rich repeat protein

SQ Sequence 259 AA;

Query Match 75.4%; Score 1315; DB 5; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MYVLSPEVFIILQLLFIQAISSSLKGFSLAMELAHRCNVDPVSTLTPVKTSEFENFKT 60
 1 MYVLSPEVFIILQLLFIQAISSSLKGFSLAMELAHRCNVDPVSTLTPVKTSEFENFKT 60
 61 KQVITSKDDYPLSKNFPYSLEHLOTSYGLVRVDMRMCLKSLRKLDSLHNHKKLPATI 120
 61 KQVITSKDDYPLSKNFPYSLEHLOTSYGLVRVDMRMCLKSLRKLDSLHNHKKLPATI 120
 121 GDLHLQELNLDNHLESFVALCHSTLQKSLRSLDLKSNKIKALPVQFCOLOELKNLKL 180

Db 121 GDLHLQELNLDNHLESFVALCHSTLQKSLRSLDLKSNKIKALPVQFCOLOELKNLKL 180
 QY 181 DONELIQPKKIGQLINLRFLSARNKLPPLPSEFRNLSLEYLDLFGNTFQPKVLPVVK 240
 Db 181 DONELIQPKKIGQLINLRFLSARNKLPPLPSEFRNLSLEYLDLFGNTFQPKVLPVVK 240
 QY 241 LQAPLTILESSARTILNHR 259
 Db 241 LQAPLTILESSARTILNHR 259
 RESULT 7
 AAU20511
 ID AAU20511 standard; protein; 259 AA.
 AC AAU20511;
 XX
 XX 06-DEC-2001 (first entry)
 XX Human secreted protein, Seq ID No 503.
 DE
 DE Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
 KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
 KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
 KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
 KW Gaucher's disease; neurological disease; cerebrovascular disorder;
 KW thrombosis; wound healing.
 XX
 XX Homo sapiens.
 XX
 XX WO200155326-A2.
 PN
 PN 02-AUG-2001.
 PD
 PD 17-JAN-2001; 2001WO-US001347.
 PF
 PF 31-JAN-2000; 2000US-017906SP.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-022668-P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 03-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249399P.
 PR 01-DEC-2000; 2000US-0250300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2000US-0254097P.
 PR 05-JAN-2001; 2000US-02559678P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI MPI; 2001-451931/48.
 XX N-PSDB; AAS33220.

New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.

Claim 11; SEQ ID NO 503; 753pp; English.

The invention relates to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and thrombosis), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). (I) and (II), agonists, antagonists and antibodies can also be used to promote wound healing, maintain organs before transplantation, and support cell culture of primary tissues.

Query Match 65.2%; Score 1137; DB 4; Length 259;
 Best Local Similarity 98.7%; Pred. No. 3.8e-107;
 Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 QAISSSLKGLSAMLRAHRCNVDPVSLTPVKCTSEFENFKMTWTSKDYPLSKNFP 77

Db	:	33	KAXSSSLXGFLSAMRLAHRGCVDPVSTLTPVKSEFENFKTKWITSKKOYPLSKNPF	92
Qy	78	YSLEHLQTSYCGELVRDWMRLCLSLRKLIDLSHRHKKLPATIGDLIHQELNDNHL	137	
Db	93	YSLEHLQTSYCGELVRDWMRLCLSLRKLIDLSHRHKKLPATIGDLIHQELNDNHL	152	
Qy	138	SFSVALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELKNLKDDELQFPCKTGOLIN	197	
Db	153	SFSVALCHSTLQKSLXSLDLSKNKIKALPVQFCQLQELKNLKDDELQFPCKTGOLIN	212	
Qy	198	LRFISAANKLPFPSPSEPNLSLEYDLFGNFTFEQPKLVPVTKLQ	242	
Db	213	LRFISAANKLPFPSPSEPNLSLEYDLFGNFTFEQPKLVPVTKLQ	257	

RESULT 8	
ABB62072	
ID	ABB62072 standard; protein; 428 AA.
XX	
AC	ABB62072;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 13008.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.

XX	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
PT	
PT	
PT	
XX	
XX	Disclosure; SEQ ID NO 13008; 21pp + Sequence Listing; English.
XX	
XX	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA sequences (ABU1840-ABU16175) and the encoded proteins (ABE57737-ABE77072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 428 AA;

Qy	70	YPLSKNPPYSLEH:QTSYCVGLVYVDMRMCLKSLRKLDLSHNHIKKLPATIGDLIHLQEL	129
Db	147	YPI-KGPPERTIKS:TTNNSQLVKSFEICTTUNTKLDVSGNKLKTIKPISELGRL-PLTSL	204
Qy	130	MLNDNHLSEFS:-----VALCHSTLOKSLRSLDSKNKI KALPVQFCQLOQLKKNLKDD	182
Db	205	HLGNLLGTQNDWCMLRCKLC-----QSLGELDLSGNGLYFPFPIVKFSELSVLNN	259
Qy	183	NELIOFPCKIGQLNLRPLSARNKLP:PFSEFRNLSLEYLDLFONTPE-----OPKV	235
Db	260	NLLSRLPFAIRRMKALRKLKYCSNELESPLSAVEDLRIDLLDVMGNCPEFNADAQQMY	319
Qy	236	LPVTKLOAPITLLESSARTILENIRIPYGSHTIIPHLQDLDTAKTICVGRFCL	288
Db	320	LQKAANSQPQMLLGRANDVKYMLPISAGSIPAVLIDIREAPRCPCGELCYAQRKEDL	379
Qy	289	-----NSFIQSTTMNL-----HSVAHTVVVLVDN	312
Db	380	FORVQPKFT---TVRNLTYSREHQIVADVVLCDS	411

RESULT 9	
AB014780	
ID	AB014780 standard; protein; 389 AA.
XX	
XX	AB014780;
AC	
XX	
XX	25-AUG-2003 (first entry)
DT	
DT	Novel human protein #153.
DE	
DE	
XX	
XX	Human; NOV; Gene therapy; endocrine related disease; diabetes;
KW	metabolism-related disease; obesity; central nervous system disorder;
KW	Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW	schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW	psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW	inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW	colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW	prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW	lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW	stroke; infection.

Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
 PI
 Gerlach VL, Vernet CAM, Ellerman K, Berchtz C, Rothenberg ME, Guo X;
 PI
 Shinkens RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
 PI
 Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
 PI

[illegible]

KW	prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW	lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW	stroke; infection.
XX	
OS	Homo sapiens.
XX	
XX	WO2003023002-A2.
XX	
PD	20-MAR-2003.
XX	
XX	09-SEP-2002; 2002WO-US028539.
PF	
XX	07-SEP-2001; 2001US-0318120P.
PR	
XX	07-SEP-2001; 2001US-0318130P.
PR	
XX	10-SEP-2001; 2001US-0318430P.
PR	
XX	17-SEP-2001; 2001US-0322636P.
PR	
XX	17-SEP-2001; 2001US-0322781P.
PR	
XX	17-SEP-2001; 2001US-0322816P.
PR	
XX	19-SEP-2001; 2001US-0323519P.
PR	
XX	20-SEP-2001; 2001US-0323631P.
PR	
XX	20-SEP-2001; 2001US-0323636P.
PR	
XX	25-SEP-2001; 2001US-0324969P.
PR	
XX	25-SEP-2001; 2001US-0325091P.
PR	
XX	26-SEP-2001; 2001US-0324990P.
PR	
XX	17-APR-2002; 2002US-0373212P.
PR	
XX	06-SEP-2002; 2002US-00236177.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI	Gierlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI	Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI	Rieger DK, Taupier RJ, Shenoy SG, Shenoy M, Alsobrook JP;
PI	Lepley DM, Edinger SE, Burgess CE;
XX	
XX	WPI; 2003-313242/30.
DR	
DR	N-PSDB; ACD19472.
XX	
XX	New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT	and polynucleotides, useful in gene therapy, e.g. for treating or
PT	preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT	stroke or infections.
XX	
XX	Claim 1; Page 411; 586pp; English.
PS	
XX	
CC	The invention describes a new isolated polypeptide (NOVX). The NOVX
CC	polypeptide, nucleic acid and antibody are useful as therapeutics,
CC	particularly in the manufacture of a medicament for treating a syndrome
CC	associated with a human disease, which includes a pathology associated
CC	with NOVX polypeptide, the DNA encoding the protein is useful in gene
CC	therapy for treating the disease or condition. In particular, the NOVX
CC	polypeptide or polynucleotide is useful for treating endocrine/
CC	metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC	system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC	epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC	and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC	asthma, inflammatory bowel disease, rheumatoid arthritis or
CC	osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC	prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC	cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC	haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC	These are also useful in developing powerful assay system for functional
CC	analysis of various human disorders, as well as in diagnostic
CC	applications, and for monitoring the effects of drugs during clinical
CC	trials. This is the amino acid sequence of a novel human NOV protein
XX	
SQ	Sequence 1489 AA;

Query Match 12.4%; Score 2-7; DB 6; Length 1489;
Best Local Similarity 28.2%; Pred. No. 3e-12;
Matches 30; Coverage 100%;

Qy	75	NPFPYSLHQTQSYCGLVVRVDMRMLCLKSURKLDLSNNHHKKKLPATIGDGLIHHQELNLNDN	133
Dd	43	NPFTLEELYLDAQITEELPKQLFNQCQALRKLSIFDNDLSNLPPTTASTASVNLKELDISEKN	102
Qy	135	HLESEF--SVALCHSTLQKSLRSLDLSKKNIKALPVQFCQLOELKKNLKLDNDELIOFPCKI	192
Dd	103	GVQEPENIKKC-----KCLTIIASVNPISKLPDGFOTLLNLTOIYNKDAFLFELPANF	157
Qy	193	GOLINRFLSAARNKLPFLPSEFRNLSELYLDLFGNTP--EQPKVLVPVKLQAPLTLLESS	251
Dd	158	GRIVAKRLRIELRENNHLKTLPKMHKLAQLERLDGNNSEFSELPVLDQIQNLREL--WMDNN	216
Qy	252	ARTILHNRIPIYGSHILPHELQOD-LDTAKICVCFRCFLNSPFOGTTMNLHVSVAIVTVJ	310
Dd	217	ALQVLPGSGIKMLVLYLDMSKNRIETVMDMDISGCEALDLSNNMLQ-----QLP	268
Qy	311	DNLGG 315	
Dd	269	DSIGG 273	
RESULT 11			
ABU11038	ID	ABU11038 standard; protein; 1489 AA.	
XX	AC	ABU11038;	
XX	AC		
XX	DT	05-FEB-2003 (first entry)	
XX	DE	Human protein NOV7.	
Kw	Kw	Human; NOVX; adrenoleukodystrophy; haemophilia; stoke; VHL;	
Kw	Kw	congenital adrenal hyperplasia; haemophilia; hypercoagulation;	
Kw	Kw	idiopathic thrombocytopaenic purpura; autoimmune disease; allergy;	
Kw	Kw	immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;	
Kw	Kw	Alzheimer's disease; tuberosus sclerosis; Parkinson's disease; epilepsy;	
Kw	Kw	Huntington's disease; cerebral palsy; Lesch-Nyhan syndrome; pain;	
Kw	Kw	multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;	
Kw	Kw	behavioural disorder; addiction; neuroprotection; diabetes; ARDS;	
Kw	Kw	renal artery stenosis; interstitial nephritis; glomerulonephritis;	
Kw	Kw	polycystic kidney disease; systemic lupus erythematosus; IGA;	
Kw	Kw	renal tubular acidosis; immunoglobulin A nephropathy; hypercalcaemia;	
Kw	Kw	cirrhosis; transplantation; asthma; emphysema; scleroderma; GVHD;	
Kw	Kw	adult respiratory distress syndrome; graft versus host disease;	
Kw	Kw	lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;	
Kw	Kw	anaemia; cancer; trauma; regeneration; infection.	
XX	OS	Homo sapiens.	
XX	OS		
XX	PN	WC200281629-A2.	
XX	PD		
XX	PD	17-OCT-2002.	
XX	PF		
XX	PF	03-APR-2002; 2002WO-US010522.	
XX	PR		
XX	PR	03-APR-2001; 2001US-0281086P.	
XX	PR	03-APR-2001; 2001US-0281136P.	
XX	PR	05-APR-2001; 2001US-0281863P.	
XX	PR	05-APR-2001; 2001US-0281906P.	
XX	PR	06-APR-2001; 2001US-0282020P.	
XX	PR	10-APR-2001; 2001US-0282934P.	
XX	PR	12-APR-2001; 2001US-0283512P.	
XX	PR	13-APR-2001; 2001US-0285325P.	
XX	PR	23-APR-2001; 2001US-0285890P.	
XX	PR	24-APR-2001; 2001US-0286068P.	
XX	PR	25-APR-2001; 2001US-0286292P.	
XX	PR	27-APR-2001; 2001US-0287213P.	
XX	PR	02-MAY-2001; 2001US-0288257P.	
XX	PR	12-MAY-2001; 2001US-0291134P.	
XX	PR	17-MAY-2001; 2001US-0291725P.	
XX	PR	31-MAY-2001; 2001US-0294771P.	
XX	PR	08-JUN-2001; 2001US-0296965P.	
XX	PR	18-JUN-2001; 2001US-0299128P.	

QY 252 ARTLHNRIPYGVSHLIPHLQD-LDTAKICVCGRFCLNSFIQGTTHLHSHVAHTVIV 310
 Db 217 ALQVLPSIGKRLMLVYDMSKNRIETVDMISGCEALEDLSSNMLQ-----QLP 268
 QY 311 DNLGG 315
 Db 269 DSIQG 273

RESULT 12
 ABB61551
 ID ABB61551 standard; protein; 431 AA.
 XX
 AC ABB61551;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 11445.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL05654.
 XX
 FT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 11445; 2lpp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 431 AA;
 Query Match 12.1%; Score 211.5; DB 4; Length 431;
 Best Local Similarity 31.6%; Pred. No. 1.8e-12;
 Matches 79; Conservative 39; Mismatches 105; Indels 27; Gaps 11;

QY 12 LQLEFQA--ISSSLKGFLSAMLAHRCNVDPFVSTLTPVKTSEFENFKTKMVI-TSKK 68
 Db 39 JEEFLFDAMHIRDLPKNFFLHRLKLGSL-DNEIGRLPP-----DIQNFENLVELDVRN 93
 QY 69 DYPLSKNFFYSLEHLOT-----SYCGLVVDVDMRLCLSLKSLKLDLSHNHKKLPATIGD 122
 Db 94 DIP---DIPDDXHLQSLQVADESSNPIKLPSPGSQLKMLTVIGLNDMSLTLPADFGS 150
 QY 123 LHLQELNLDNHELESFVALCHSTLQKLSRLSDLSKNIKALPVQFCQLQELKMLKLD 182

Db 151 LTQLESLELRNLLKHLPETISQLT---KLKRLDLDGNEIEDLPYLGYPGLHMLWDH 207
 QY 183 NELIQFCKIGQLINRLSAAARKLPPLPSEFRL-SLEYLDLFGNTFEQPKVLP--VI 239
 Db 208 NQLQELPELGLLTKLYLDVSENRLSELPNEISGLVSTLDDLAQNLLS---ALPDGIA 264
 QY 240 KLOAPLTLLE 249
 Db 265 KL-SRLTILK 273

RESULT 13
 ABB63454
 ID ABB63454 standard; protein; 378 AA.
 XX
 AC ABB63454;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 23154.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL09557.
 XX
 FT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 23154; 2lpp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 378 AA;
 Query Match 11.9%; Score 207.5; DB 4; Length 378;
 Best Local Similarity 32.0%; Pred. No. 3.8e-12;
 Matches 66; Conservative 39; Mismatches 72; Indels 29; Gaps 8;

QY 82 HLOTSVCGLVRVDMRLCLSLKSLKLDLSHNHKKLPATIGLHQLQELNLDNHELESFV 141
 Db 53 HLAGNLSLPELDVIY--LENLEFLDVSNNEKELPPTLGLLNLQQLNYSNGNLTELPV 110
 QY 142 ALCHSTLQKLSRLSDLSKNIKALPVQFCQLQELKMLKLDNE-LIQFCKIGQLINLR 200
 Db 111 EL--SGL--RNLEHNLGNKQCFRALPVQLSECVRLNEVNSDNEALVHIPERISNLPMLQS 167

QY 201 LSAARNKLPFSE-----FRNLSLEYLDLFGNTFEQ-----PKVLPVTKQAP 244
 DB 168 LAADRCALVTLPAALSKEFANHRIFHTAINIIPMYVERFYQNFYDRQRTPTVAVERKG 227
 QY 245 L---TLESARTILHNRIPVGHII 267
 DB 228 LFWVRELETSTRLLL-----PVGRTV 249

RESULT 14

ABB64971

ID ABB64971 standard; protein; 849 AA.

XX AC ABB64971;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 21705.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX OS pharmaceutical.

XX PN Drosophila melanogaster.

XX PR WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL09074.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

XX PT genes from Drosophila and for elucidating cell signalling and cell-cell

XX PT interactions.

XX PS Disclosure; SEQ ID NO 21705; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

XX CC capable of detecting 1000 or more genes from Drosophila. The invention is

XX CC useful in developmental biology and in elucidating cell signalling and

XX CC cell-cell interactions in higher eukaryotes for the development of

XX CC insecticides, therapeutics and pharmaceutical drugs. The invention

XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

XX CC ABB72072). The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC Sequence 849 AA;

XX CC Query Match 11.9%; Score 207; DB 4; Length 849;

XX CC Best Local Similarity 33.5%; Pred No. 1.4e-11;

XX CC Matches 69; Conservative 34; Mismatches 83; Indels 20; Gaps 8;

QY 49 PVKTSFNFNFKTM-VITSKDY---PLSKNFP-----YSLHQLQTSYCGVLRVDMRML 98

DB 2 PLLSKCFPCFKREVIDKLDYSNTELT-DPEVWQHURTEELVLTTR-QALPPQLF 60

QY 99 CLKSLRLKLDLSHNHKKLPATIGDLHQLNLDNHLSP--SYVALCHSTQKSLRLSD 156

DB 61 YCQGLRVLVNHNLESIPQAGISLRQQLQHLNENLVNVPPEIKSC-----KEFTHLD 115

QY 157 LSKNKIKALPVQFCQLOELKNLKLDDNELIOFPCKIGQLINRFLSARNKLPFLP-SEF 215

DB 157 LSKNKIKALPVQFCQLOELKNLKLDDNELIOFPCKIGQLINRFLSARNKLPFLP-SEF 215

DB 116 LSCNSLQRLPDATSLISLQELINETYELFPAFGRLVNLRIELRLNNLMTLPKSNV 175
 QY 216 RNLSLEYLDLFGNTFEQ-BQPKVLPVTK 240
 DB 176 RLINLQRLDIDGNEFTLPELVVGELK 201

RESULT 15

ABB67403

ID ABB67403 standard; protein; 1060 AA.

XX AC ABB67403;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 29001.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX OS pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL11506.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

XX PT genes from Drosophila and for elucidating cell signalling and cell-cell

XX PT interactions.

XX PS Disclosure; SEQ ID NO 29001; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

XX CC capable of detecting 1000 or more genes from Drosophila. The invention is

XX CC useful in developmental biology and in elucidating cell signalling and

XX CC cell-cell interactions in higher eukaryotes for the development of

XX CC insecticides, therapeutics and pharmaceutical drugs. The invention

XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

XX CC ABB72072). The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC Sequence 1060 AA;

XX CC Query Match 11.5%; Score 201.5; DB 4; Length 1060;

XX CC Best Local Similarity 33.7%; Pred. No. 6.9e-11;

XX CC Matches 63; Conservative 30; Mismatches 63; Indels 31; Gaps 6;

QY 75 NFFYSLEHLQTSYCGVLRVDMRMLCLKSLRKLDLSHNHKKLPATIGDLHQLNLDN 134

DB 204 NFFTSIDSLAN-----LCLSHNSLPLKLPDCVYVTVTLVRLNLSN 246

QY 135 HUESFVALCHSTQKSLRLSDLSKVKALPVQFCQLOELKNLKLDDNEL--IOFPCKI 192

DB 247 ETELTAGV---ELWQRLSINLSRNQVLPALCKLPCKLRLLVNDKLNPEGIPSGI 303

QY 193 GQLINRFLSARNKLPFLPSEF-RNLS-EYLDLFGNTFEQKVPVTKQAP-LLES 251

DB 304 KXGALGVFSAANNLEWVPEGLRCGALKQLNLSNR-----L-LTPDA-LHLEGL 355

Wed Jun 16 15:55:24 2004

us-10-009-557-9.rag

Page 12

Qy 252 ARTILRN 258
Db 356 DQLDLRN 362

Search completed: June 15, 2004, 10:01:19
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 22:58:17 ; Search time 2398 Seconds
(without alignments)
15130.338 Million cell updates/sec

Title: US-10-009-557-34

Perfect score: 1215

Sequence: 1 gaagaactagcatgtatgta.....ttagtttaaaaaaaaaa 1215

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_estum.*

3: em_estin.*

4: em_estu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_estl.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfn.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_man.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_pkg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1091.2	89.8	1437	11	CNSLTI186
2	1050	86.4	1143	9	AL582092
3	956.6	78.7	989	9	AL582114
4	925.8	76.2	1155	9	AL560571

C	5	902.4	74.3	1133	13	BX405954
C	6	764.8	62.9	770	13	BX401074
	7	760.4	62.6	1150	13	BX333340
	8	759.2	62.5	948	13	BX370934
	9	758	62.4	791	13	BX370934
	10	733.6	60.4	901	13	BX370934
C	11	733.6	60.4	901	13	BX370934
C	12	705.2	58.0	1049	13	BX370934
C	13	703.8	57.9	722	13	BX370934
C	14	700	57.6	891	13	BX370934
C	15	695.8	57.3	1165	13	BX370934
C	16	684	56.3	947	13	BX370934
C	17	660.4	54.4	972	13	BX370934
C	18	649.6	53.5	1201	13	BX370934
C	19	646.8	53.2	991	9	AL560705
C	20	627.6	51.7	1032	12	SM554257
C	21	610.4	50.2	1157	13	BX405955
C	22	582.8	48.0	1072	12	BM476830
C	23	538	44.3	613	9	AV717878
C	24	501.2	41.3	685	29	CE788486
C	25	479.2	39.4	658	14	CF953469
C	26	476.6	39.2	572	9	AI739536
C	27	462.6	38.1	1050	13	BX346615
C	28	458	37.7	483	12	BG940337
C	29	449.8	37.0	874	14	CF265122
C	30	443.8	36.5	467	9	AA424400
C	31	433.8	35.7	466	9	AA812029
C	32	430.8	35.5	446	9	AA393227
C	33	430.8	35.5	645	14	CD365392
C	34	430	35.4	439	13	BU753893
C	35	430	35.4	561	14	CA421227
C	36	428.4	35.3	473	9	AA742244
C	37	422.4	34.8	775	13	BQ053807
C	38	420.8	34.6	540	10	AM971618
C	39	419	34.5	656	13	BU620172
C	40	412.2	33.9	801	12	B1258030
C	41	411.8	33.9	725	10	BG067404
C	42	411.4	33.9	999	13	BU526707
C	43	411	33.8	411	9	AA424474
C	44	409	33.7	512	9	AA648933
C	45	406.2	33.4	1002	12	BM925426

ALIGNMENTS

RESULT 1

CNSLTI186

LOCUS

DEFINITION

human full-length cDNA clone CS0DK007YE23 of HeLa cells of Homo sapiens (human).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNSLTI186 1437 bp mRNA linear HTC 18-JUN-2003
human full-length cDNA clone CS0DK007YE23 of HeLa cells of Homo sapiens (human).

BX248298.1 GI:28207942

HTC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1437)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Paradise Avenue

2 (bases 1 to 1437)

Genoscope.

Direct Submission

Submitted (04-FEB-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 86.4%; Score 1050; DB 9; Length 1143;
Best Local Similarity 97.3%; Pred. No. 4.6e-231;
Matches 1084; Conservative 13; Mismatches 13; Indels 4; Gaps 3;

QY 59 TATTCAGGCCATTCACGAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATAG 118
DB 1110 TAAGDAAGCCATTCACGAGTTTAAAGGTTT-CCTTTCATATGAACTGGCTCATAG 1053

QY 119 AGGCTGTATGTGTATACACCACTTTCAACGCTCACACGAGTGAAGACTTCAGAAATTGA 178
DB 1052 AGGCTGTAAATGKRTACACCACTTTCAACGCTCACACGAGTGAAGACTTCAGAAATTGA 993

QY 179 AAACCTTTAAACTAAATGGTTATACATCCAAAAGACTATCTCTTAAGTAAGAAATT 238
DB 992 AAACCTTTAAACTAAATGGTTATACATCCAAAAGACTATCTCTTAAGTAAGAAATT 933

QY 239 TCCATATTCCTTGGAACTTTTCAAGCTCTTACTGTGGCTTGTCCGAGTTGATGCG 298
DB 932 CCAATATTCCTTGGAACTTTTCAAGCTCTTACTGTGGCTTGTCCGAGTTGATGCG 873

QY 299 TATGCTTTGCTTAAAGCCCTTAGGAATATAGACTTGATGTCACCAACCATATAAAAGCT 358
DB 872 TATGCTTTGCTTAAAGCCCTTAGGAATATAGACTTGATGTCACCAACCATATAAAAGCT 813

QY 359 TCCAGCTACAAATGGAGACCTTATACCTTCAAGAACTTAACTGATGATGATGATG 418
DB 812 TCCAGCTACAAATGGAGACCTTATACCTTCAAGAACTTAACTGATGATGATGATG 753

QY 419 GGAGTCAATTTAGTGTAGCTTGTGTCATCTTACACTCCAGAGTCACTTCGGAGTTTGA 478
DB 752 GGAGTCAATTTAGTGTAGCTTGTGTCATCTTACACTCCAGAGTCACTTCGGAGTTTGA 693

QY 479 CCTCAGCAAGAAACAAATCAAGGCACTCCCTGTGCACTTTTGCAGCTCCAGGAACCTAA 538
DB 692 CCTCAGCAAGAAACAAATCAAGGCACTCCCTGTGCACTTTTGCAGCTCCAGGAACCTAA 633

QY 539 GAATTTAAACTTCAGCAATATGATGATCAATTCATCTTCCAGAGTAGGACACTAAT 598
DB 632 GAATTTAAACTTCAGCAATATGATGATCAATTCATCTTCCAGAGTAGGACACTAAT 573

QY 599 AAACCTTGGCTTTTGTGAGCACTCGAAATAAGCTTCCATTTTTCCTAGTGAATTAG 658
DB 572 AAACCTTGGCTTTTGTGAGCACTCGAAATAAGCTTCCATTTTTCCTAGTGAATTAG 513

QY 559 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAATCT 718
DB 512 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAATCT 453

QY 719 TCCAGTAATAAGCTGGAAGCAACCAATTAACCTTTATGGAATCTTCTGCAAGCAACATAT 778
DB 452 TCCAGTAATAAGCTGGAAGCAACCAATTAACCTTTATGGAATCTTCTGCAAGCAACATAT 393

QY 779 ACATAATAGGATTCATATGCTCTCATATCAATTCATTCATTCCTGCCAAGATTTGGA 838
DB 392 ACATAATAGGATTCATATGCTCTCATATCAATTCATTCATTCCTGCCAAGATTTGGA 333

QY 839 TACCGCAAAATTTGTTTGTGGAGATCTTCTGCAAGCTTCTTCAATTCATTCAGGAACCTAC 898
DB 332 TACCGCAAAATTTGTTTGTGGAGATCTTCTGCAAGCTTCTTCAATTCATTCAGGAACCTAC 273

QY 899 TACCATGAATCTGATCTCTGTTGCAACCTGCTGCTTCTAGTATATTTTGGTGGTAC 958
DB 272 TACCATGAATCTGATCTCTGTTGCAACCTGCTGCTTCTAGTATATTTTGGTGGTAC 214

QY 959 TGAAGCACTTATATCTCTTATTTCTGTTCTCTAGGCTGTTATGTTAATCTCTGATAT 1018
DB 213 TGAAGCACTTATATCTCTTATTTCTGTTCTCTAGGCTGTTATGTTAATCTCTGATAT 154

QY 1019 GTTAAAGTAATGGTGAGACCCAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGA 1078
DB 153 GTTAAAGTAATGGTGAGACCCAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGA 94

QY 1079 TGTATGATTTTCCAGCGTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGGAG 1138
DB 93 TGTATGATTTTCCAGCGTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGGAG 34

QY 1139 GAGGATGTGTATAGTACTTCAATTTAGATGACTC 1172
DB 33 GAGGATGTGTATAG-TACTCAATTTAGATGACTC 1

RESULT 3

AL582114/c
LOCUS AL582114 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL004YC22 3-PRIME, mRNA sequence.
ACCESSION AL582114
VERSION AL582114.2 GI:31320341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 989)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12949780.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL004B11NP1&cluster=7624.f. Contact :
Feng Jiang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL004B11NP1.
Location/Qualifiers
1 .989

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL004YC22"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 78.7%; Score 956.6; DB 9; Length 989;
Best Local Similarity 98.7%; Pred. No. 1.4e-203;
Matches 976; Conservative 8; Mismatches 3; Indels 2; Gaps 2;

QY 179 AAATTTAAACTTAAATGGTTATACATCCAAAAGACTATCTCTAAGTAAGAAATTT 238
DB 989 AAATTTAAACTTAAATGGTTT-TCACATCCAAAAGACTATCTCTAAGTAAGAAATTT 931

QY 239 TCCATATTCCTTGGAACTTTTCAAGCTTCTTACTGTGGCTTGTCCAGTTGATGCG 298
DB 930 TCCATATTCCTTGGAACTTTTCAAGCTTCTTACTGTGGCTTGTCCAGTTGATGCG 871

QY 299 TATGCTTTGCTTAAAGCCCTTAGGAATTAAGCTTGAATCAACCACTAT-AAAAAGC 357
DB 213 TATGCTTTGCTTAAAGCCCTTAGGAATTAAGCTTGAATCAACCACTAT-AAAAAGC 154

```
Db      870  TATGCTTTGCTTAAAGAGCCCTTAGGAAATTAGACTTGAGTCACACCAATATAAAAAAGC 811
Qy      358  TTCCAGCTACAAATTGGAGACCTCTACACCTCTCAAGAACTTAACTGAAATGACATCACT 417
Db      810  TTCCAGCTACAAATTGGAGACCTCTACACCTCTCAAGAACTTAACTGAAATGACATCACT 751
Qy      418  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAAAGTCACCTCGGAGTTGG 477
Db      750  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAAAGTCACCTCGGAGTTGG 691
Qy      478  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGTCAGTTTTCAGAGTCCTCCAGGAACCTTA 537
Db      690  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGTCAGTTTTCAGAGTCCTCCAGGAACCTTA 631
Qy      538  AGAATTTTAAATCTGACGATTAATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 597
Db      630  AGAATTTTAAATCTGACGATTAATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 571
Qy      598  TAAACCTTTGCTTTTGTGTCAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTA 657
Db      570  TAAACCTTTGCTTTTGTGTCAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTA 511
Qy      658  GAAATTTATCCCTTGATACCTTGATCTTTTGGAAATACCTTTTGAACAACCAAAAGTCC 717
Db      510  GAAATTTATCCCTTGATACCTTGATCTTTTGGAAATACCTTTTGAACAACCAAAAGTCC 451
Qy      718  TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTATTGGAATCTTCTGCAAGCACTAT 777
Db      450  TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTATTGGAATCTTCTGCAAGCACTAT 391
Qy      778  TACATAATAGCAATTCATATGCTCTCATATCATTCATTCATTCATCTCTGCAAGATTTGG 837
Db      390  TACATAATAGCAATTCATATGCTCTCATATCATTCATTCATTCATCTCTGCAAGATTTGG 331
Qy      838  ATACCGGAAATTTGCTTTGTGGAAGATCTGTCGAACTTCTTCAATTCATTCAGGAACCTA 897
Db      330  ATACCGGAAATTTGCTTTGTGGAAGATCTGTCGAACTTCTTCAATTCATTCAGGAACCTA 271
Qy      898  CTACATGAATCTGATCTCTGTCGCCACACTGTGCTTGTAGTAGATAATTTGGGTGGTA 957
Db      270  CTACATGAATCTGATCTCTGTCGCCACACTGTGCTTGTAGTAGATAATTTGGGTGGTA 211
Qy      958  CTGAAGCACTTATATCTCTTATTTCTCTTCTAGGCTGTTAATTTCTCTGATA 1017
Db      210  CTGAAGCACTTATATCTCTTATTTCTCTTCTAGGCTGTTAATTTCTCTGATA 151
Qy      1018  TGTAAAGTAATGGTGAGACCAAGAAAGAAATTTCAATTAACAGATCAGTTTGGGTGC 1077
Db      150  TGTAAAGTAATGGTGAGACCAAGAAAGAAATTTCAATTAACAGATCAGTTTGGGTGC 91
Qy      1078  ATGTATGATTTTGCAGGCTCAAAATGGAGTAGGGAAGATTTCTGTATACCTTGTCTGGAGA 1137
Db      90  ATGTATGATTTTGCAGGCTCAAAATGGAGTAGGGAAGATTTCTGTATACCTTGTCTGGAGA 31
Qy      1138  GGAGGAATGTGTATAGTTACTCAATTTAGA 1166
Db      30  GGAGGAATGTGTATAGTTACTCAATTTAGA 2
```

```
RESULT 4
LOCUS   AL560571
DEFINITION AL560571 Homo sapiens B CELLS (RAMOS CELL LINE) EST 31-MAY-2003
Homo sapiens cDNA clone CS0DL003VP23 5-PRIME, mRNA sequence.
ACCESSION AL560571
VERSION   AL560571.2 GI:31284701
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS  Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
```

TITLE
JOURNAL
COMMENT

Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12907161.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL003CC12QPl&cluster=7624.f. Contact :
Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL003CC12QPl.

FEATURES
source

1..1155
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL003VP23"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE); COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 76.2%; Score 925.8; DB 9; Length 1155;
Best Local Similarity 96.0%; Pred. No. 1.7e-202;
Matches 982; Conservative 17; Mismatches 19; Indels 5; Gaps 5;
Qy 58 TTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 117
Db 124 TAAGTAAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 183
Qy 118 GAGGCTGTAATGTTGATACACCGTTTCAACGCTCACACCACTGAAGACTTCAGAAATTG 177
Db 194 GAGGCTGTAATGTTGATACACCGTTTCAACGCTCACACCACTGAAGACTTCAGAAATTG 243
Qy 178 AAAACTTTAAAGCTAAATGGTTATCACATCCAAAAGAGACTATCCTCTAAGTAAGAATT 237
Db 244 AAAACTTTAAAGCTAAATGGTTATCACATCCAAAAGAGACTATCCTCTAAGTAAGAATT 303
Qy 238 TTCCATATTCCTTGGACACTCTTCAGACTTCTTACTGTGGGCTTCTCCGAGTTGATATGC 297
Db 304 TTCCATATTCCTTGGACACTCTTCAGACTTCTTACTGTGGGCTTCTCCGAGTTGATATGC 363
Qy 298 GTATGCTTTGCTTAAAGAGCCCTTAGGAAATTAGACTTCAGTTCACACCACTATAAAAAAGC 357
Db 364 GTATGCTTTGCTTAAAGAGCCCTTAGGAAATTAGACTTCAGTTCACACCACTATAAAAAAGC 423
Qy 358 TTCCAGCTACAAATTGGAGACCTCTATACACCTTCAAGAACTTAACTGAAATGACATCACT 417
Db 424 TTCCAGCTACAAATTGGAGACCTCTATACACCTTCAAGAACTTAACTGAAATGACATCACT 483
Qy 418 TGGAGTCATTTAGTGTAGCTTGTCTATCTTACACTCCAGAAAGTCACCTCGGAGTTTGG 477
Db 484 TGGAGTCATTTAGTGTAGCTTGTCTATCTTACACTCCAGAAAGTCACCTCGGAGTTTGG 543
Qy 478 ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGTCAGTTTTCAGAGTCCTCCAGGAACCTTA 537
Db 544 ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGTCAGTTTTCAGAGTCCTCCAGGAACCTTA 603
Qy 538 AGAATTTTAAATCTGACGATTAATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 597
Db 604 AGAATTTTAAATCTGACGATTAATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 663
Qy 598 TAAACCTTTGCTTTTGTGTCAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTA 657

Db 664 TAAACCTTTCGCTTTTGTGTCAGCAGCTCGAAATAGCTTCCATTTTGTGCTAGTGAATTTA 723
 Qy 658 GAAATTTATCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAACCAAAAGTCC 717
 Db 724 GAAATTTATCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAACCAAAAGTCC 783
 Qy 718 TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTTATTTGGAATCTTCTGCAAGCAACATAT 777
 Db 784 TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTTATTTGGAATCTTCTGCAAGCAACATAT 843
 Qy 778 TACATAATAGATTCCTATGCTCTCATATCTTCCATTCATCTTCCATTCCTGCGCAAGATTTGG 837
 Db 844 TACATAATAGATTCCTATGCTCTCATATCTTCCATTCATCTTCCATTCCTGCGCAAGATTTGG 903
 Qy 838 ATACCGCAAAATTTGT-TTTTGTGGAAGATCTCTCTGAACTCTTTTCATT-CAAGGAAC 895
 Db 904 ATACCGCAAAATTTGTGGTTTGTGGAAGATCTCTCTGAACTCTTTTCATT-CAAGGAAC 963
 Qy 896 TACTACCATGAATCTG-CATTCTGTGTGCGCCACATCTGTCTTACTAGATATTTGGGTG 954
 Db 964 TACTACCATGAATCTGCAATCTGTGTGCGCCACATCTGTCTTACTAGATATTTGGGTG 1023
 Qy 955 GTACTGAGCACCATTATCTC-TTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCT 1013
 Db 1024 GTACTGAGCACCATTATCTC-TTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCT 1083
 Qy 1014 CATATGTTTAAAGTAAATGGTGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTTGG 1073
 Db 1084 CATATGTTTAAAGTAAATGGTGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTTGG 1142
 Qy 1074 GTG 1076
 Db 1143 GKG 1145

RESULT 5

LOCUS BX405954/c
 DEFINITION BX405954 Homo sapiens FETAL LIVER Homo sapiens cdna clone
 CSODM009YF13 3-PRIME, mRNA sequence.
 ACCESSION BX405954
 VERSION BX405954.1 GI:30635447
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1133)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 7624.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSOAM009CC07NP1&cluster=7624.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSOAM009CC07NP1.
 Location/Qualifiers
 1. .1133
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODM009YF13"
 /tissue type="FETAL LIVER"
 /dev stage="fetal"
 /clone lib="Homo sapiens FETAL LIVER"
 /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cdna"

FEATURES
source

Query Match 74.3%; Score 902.4; DB 13; Length 1133;
 Best Local Similarity 97.1%; Pred. No. 4.1e-197;
 Matches 941; Conservative 16; Mismatches 8; Indels 4; Gaps 4;
 ORIGIN
 Qy 180 AACTTTAAACATAAATGGTTATCATCCAAAAAGACTATCTCTTAAGTAAGATTTT 239
 Db 974 AACTTTAAATTAATGGTTAT-ACATCCAAAAAARATWCCCTCATAGTAAGAA-TTT 917
 Qy 240 CCAATATCTCTTGGACATCTTCAGACTCTTACTTGGGCTTGTCCGAGTGAATTCGCT 299
 Db 916 CCAATATCTCTTGGACATCTTCAGACTCTTACTTGGGCTTGTCCGAGTGAATTCGCT 957
 Qy 300 ATGCTTTGCTTAAAAAGCCCTT-AGGAAATTAGACTTGAGTCAACACCATATAAAAGCT 358
 Db 856 ATGCTTTGCTTAAAAAGCCCTTAAAGGAATTAGACTTGAGTCAACACCATATAAAAGCT 797
 Qy 359 TCCAGCTACAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTT 418
 Db 796 TCCAGCTACAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTT 737
 Qy 419 GGAGTCATTTAGTGTAGCTTGTGTCATCTACACTCCAGAACTCACTTCGGAGTTTGA 478
 Db 736 GGAGTCATTTAGTGTAGCTTGTGTCATCTACACTCCAGAACTCACTTCGGAGTTTGA 677
 Qy 479 CCTCAGCAAGAACAAATCAAGGCACTCCCTGTGAGTTTGCAGCTCCAGAACTTAA 538
 Db 676 CCTCAGCAAGAACAAATCAAGGCACTCCCTGTGAGTTTGCAGCTCCAGAACTTAA 617
 Qy 539 GAATTTAAACCTTGACGATATGAATTTGATTCATTTCTTGCAGATAGGCAACTAAT 598
 Db 616 GAATTTAAACCTTGACGATATGAATTTGATTCATTTCCCTTGAAGATAGGCAACTAAT 558
 Qy 599 AAACCTTCGCTTTTGTTCAGCAGCTCGAAATGAAGCTTCGATTTTTCCTAGTGAATTTAG 658
 Db 557 AAACCTTCGCTTTTGTTCAGCAGCTCGAAATGAAGCTTCGATTTTTCCTAGTGAATTTAG 498
 Qy 659 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAAACCAAAAGTCT 718
 Db 457 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAAACCAAAAGTCT 438
 Qy 718 TCCAGTAATAAGCTGCAAGCACCATTAACTTTATTTGAACTTCTGCGAAGAACCTAT 778
 Db 437 TCCAGTAATAAGCTGCAAGCACCATTAACTTTATTTGAACTTCTGCGAAGAACCTAT 378
 Qy 779 ACATAATAGGATTCATATGCTCTCATATCCATTCATTCATTCATTCATTCATTCATTCAT 838
 Db 317 ACATAATAGGATTCATATGCTCTCATATCCATTCATTCATTCATTCATTCATTCATTCAT 318
 Qy 839 TACCGCAAAATTTGTGTTTGTGGAAGATTTCTGCTGAACCTTTTTCATTCAAGGAACCTAC 898
 Db 317 TACCGCAAAATTTGTGTTTGTGGAAGATTTCTGCTGAACCTTTTTCATTCAAGGAACCTAC 258
 Qy 899 TACCATGAATCTGATTTGTGTTGCGCACACTGTGTCTTAGTAGATAATTGGGTGGTAC 958
 Db 257 TACCATGAATCTGATTTGTGTTGCGCACACTGTGTCTTAGTAGATAATTGGGTGGTAC 198
 Qy 959 TGAAGCACTATCTATCTTATTTCTGTCTTAGGCTGTTATTTAACTTCTCTCGATAT 1018
 Db 197 TGAAGCACTATCTATCTTATTTCTGTCTTAGGCTGTTATTTAACTTCTCTCGATAT 138
 Qy 1019 GTTAAAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGTGCA 1078
 Db 137 GTTAAAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGTGCA 78
 Qy 1079 TGTATGATTTTGCAGGCTCAAAATGGAGTAAGGGAAGATTTCTGTATCTTCTGCTGGAGAG 1138
 Db 77 TGTATGATTTTGCAGGCTCAAAATGGAGTAAGGGAAGATTTCTGTATCTTCTGCTGGAGAG 18

Db 664 TAAACCTTTCGCTTTTGTGTCAGCAGCTCGAAATAGCTTCCATTTTGTGCTAGTGAATTTA 723
 Qy 658 GAAATTTATCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAACCAAAAGTCC 717
 Db 724 GAAATTTATCCTTGAATCTTGGATCTTTTGGAAATACTTTTGAACAACCAAAAGTCC 783
 Qy 718 TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTTATTTGGAATCTTCTGCAAGCAACATAT 777
 Db 784 TTCCAGTAATAAAGCTGCAAGCACCATTAACTTTTATTTGGAATCTTCTGCAAGCAACATAT 843
 Qy 778 TACATAATAGATTCCTATGCTCTCATATCTTCCATTCATCTTCCATTCCTGCGCAAGATTTGG 837
 Db 844 TACATAATAGATTCCTATGCTCTCATATCTTCCATTCATCTTCCATTCCTGCGCAAGATTTGG 903
 Qy 838 ATACCGCAAAATTTGT-TTTTGTGGAAGATCTCTCTGAACTCTTTTCATT-CAAGGAAC 895
 Db 904 ATACCGCAAAATTTGTGGTTTGTGGAAGATCTCTCTGAACTCTTTTCATT-CAAGGAAC 963
 Qy 896 TACTACCATGAATCTG-CATTCTGTGTGCGCCACATCTGTCTTACTAGATATTTGGGTG 954
 Db 964 TACTACCATGAATCTGCAATCTGTGTGCGCCACATCTGTCTTACTAGATATTTGGGTG 1023
 Qy 955 GTACTGAGCACCATTATCTC-TTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCT 1013
 Db 1024 GTACTGAGCACCATTATCTC-TTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCT 1083
 Qy 1014 CATATGTTTAAAGTAAATGGTGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTTGG 1073
 Db 1084 CATATGTTTAAAGTAAATGGTGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTTGG 1142
 Qy 1074 GTG 1076
 Db 1143 GKG 1145

QY 1139 GAGGAATGT 1147
DB 17 GAGGAATKT 9

RESULT 6
LOCUS BX401074/c
DEFINITION BX401074 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION BX401074
VERSION BX401074.1 GI:30614440
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 770)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK007AC12NP1&cluster=7624.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK007AC12NP1.
Location/Qualifiers
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK007YB23"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_lines="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

Query Match 62.9%; Score 764.8; DB 13; Length 770;
Best Local Similarity 99.4%; Pred. No. 1.9e-165;
Matches 765; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 384 CACCTTCAGAACCTTAACCTGCAATGACATCACTTGGAGTCATTAGTAGCCCTTGCT 443
DB 770 CACCTTCAGAACCTTAACCTGCAATGACATCACTTGGAGTCATTAGTAGCCCTTGCT 711

QY 444 CATCTACACTCCAGAACTCACTTCGGAGTTGGACCTCAGCAAGCAACAAATCAAGCA 503
DB 710 CATCTACACTCCAGAACTCACTTCGGAGTTGGACCTCAGCAAGCAACAAATCAAGCA 651

QY 504 CTCCTGTGCGAGTTTGGCAGCTCCAGAACTTAAGAAATTAATAAATTTGACGATAATGAA 563
DB 650 CTCCTGTGCGAGTTTGGCAGCTCCAGAACTTAAGAAATTTAATAAATTTGACGATAATGAA 591

QY 564 TTGATTCGAATTCCTTCGCAAGTAGGACATCAATAAACCTTCGCTTTTGTGACGACT 623
DB 590 TTGATTCGAATTCCTTCGCAAGTAGGACATCAATAAACCTTCGCTTTTGTGACGACT 631

QY 624 CGAATAAGCTTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGAT 683
DB 530 CGAATAAGCTTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGAT 471

QY 684 CTTTGGAAATACTTTTGAACCAACCAAAAGTCTCTCCAGTAAATAGAGCTCGACGACCA 743
DB 470 CTTTGGAAATACTTTTGAACCAACCAAAAGGCTTCCAGTAAATAAGCGCAAGCACA 411

QY 744 TTAACCTTTATGGAATCTTCTGCACGAACCATATTACATAATAGGATTCATATGGCTCT 803
DB 410 TTAACCTTTATGGAATCTTCTGCACGAACCATATTACATAATAGGATTCATATGGCTCT 351

QY 804 CATATCATTCATTCATCTCTCCAGAAATTTGGATACCGCAAAAATTTCTGTTGTGGA 863
DB 350 CATATCATTCATTCATCTCTCCAGAAATTTGGATACCGCAAAAATTTCTGTTGTGGA 291

QY 864 AGATTCCTGCTGAACCTCTTTCATTCAAGGAACCTACTACCAATGAATCTGCAATCTGTTGCC 923
DB 290 AGATTCCTGCTGAACCTCTTTCATTCAAGGAACCTACTACCAATGAATCTGCAATCTGTTGCC 231

QY 924 CACACTGTGGTCTTACTAGATAATTTGGGTGGTACTGAAGCACTATTATCTCTTATTTTC 983
DB 230 CACACTGTGGTCTTACTAGATAATTTGGGTGGTACTGAAGCACTATTATCTCTTATTTTC 171

QY 984 TGTCTCTAGGCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGGTGAGACAGAA 1043
DB 170 TGTCTCTAGGCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGGTGAGACAGAA 111

QY 1044 AAAGAAATTCATTAACACAGATCAGTTTGGGTGGTACTGAAGCACTATTATCTCTTATTTTC 1103
DB 110 AAAGAAATTCATTAACACAGATCAGTTTGGGTGGTACTGAAGCACTATTATCTCTTATTTTC 51

QY 1104 GAGTAAGGAAGATTTCTGTATATCTTCTGAGAGGAGGAGGAATGTGTATAG 1153
DB 50 GAGTAAGGAAGATTTCTGTATATCTTCTGAGAGGAGGAGGAATGTGTATAG 1

RESULT 7
LOCUS BX333340
DEFINITION BX333340 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION BX333340
VERSION BX333340.1 GI:30308198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1150)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK024AB03QP2
cluster=7624.f. Contact : Feng Liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK024AB03QP2.
Location/Qualifiers
1..1150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK024YC05"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 62.6%; Score 760.4; DB 13; Length 1150;
Best Local Similarity 99.9%; Pred. No. 1.9e-164;
Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 454 TCAGAGTACCTTCGGAGTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGC 513
DB 62 TCAGAGTACCTTCGGAGTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGC 121

QY 514 AGTTTTCGAGCTCCAGGAACCTTAAGAAATTAAGAACTTGACGATATGAATTCATTCAT 573
DB 122 AGTTTTCGAGCTCCAGGAACCTTAAGAAATTAAGAACTTGACGATATGAATTCATTCAT 181

QY 574 TTCTTCGAGGATAGACCACTTAATAAAGCTTCGCTTTTGTTCAGCAGCTCGAAATAAGC 633
DB 182 TTCTTCGAGGATAGACCACTTAATAAAGCTTCGCTTTTGTTCAGCAGCTCGAAATAAGC 241

QY 634 TTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGATATCTTGATCTTTTGGAA 693
DB 242 TTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGATATCTTGATCTTTTGGAA 301

QY 694 ATACTTTTGAACCAACCAAGGCTCTCCAGTATAAAGCTGCAAGCACCATTAACTTTAT 753
DB 302 ATACTTTTGAACCAACCAAGGCTCTCCAGTATAAAGCTGCAAGCACCATTAACTTTAT 361

QY 754 TGGAACTTTCTGACGAAACCAATATATACATAATAGGATTCATATGCTCTCATATCATTC 813
DB 362 TGGAACTTTCTGACGAAACCAATATATACATAATAGGATTCATATGCTCTCATATCATTC 421

QY 814 CATTCATCTCTGCAAGATTTGGATACCGCAAAATTTGCTTGTGCAAGATCTGTGC 873
DB 422 CATTCATCTCTGCAAGATTTGGATACCGCAAAATTTGCTTGTGCAAGATCTGTGC 481

QY 874 TGAACCTTTTCAATCAAGAACTTACCATGATCTGCAATCTGTTTGGCCACACCTGTGC 933
DB 482 TGAACCTTTTCAATCAAGAACTTACCATGATCTGCAATCTGTTTGGCCACACCTGTGC 541

QY 934 TCTTAGTAGAATATTTGGTGTGCTAGAGACCTATATCTCTTATTTCTGTCTCTAG 993
DB 542 TCTTAGTAGAATATTTGGTGTGCTAGAGACCTATATCTCTTATTTCTGTCTCTAG 601

QY 994 GCTGTATGTAAATTCCTCTGATATGTTAAAGTATGGGTGAGACCAAGAAAGAAATTT 1053
DB 602 GCTGTATGTAAATTCCTCTGATATGTTAAAGTATGGGTGAGACCAAGAAAGAAATTT 661

QY 1054 CAATAACAGATCAGTTTGGGTGCAATGATGATTTTTCAGCGCTCAAAATGGAGTAAGGA 1113
DB 662 CAATAACAGATCAGTTTGGGTGCAATGATGATTTTTCAGCGCTCAAAATGGAGTAAGGA 721

QY 1114 AGATTTCTGTATCTGCTGAGAGAGCAATGCTATAGTTACTCATTTAGATCATCTCC 1173
DB 722 AGATTTCTGTATCTGCTGAGAGAGCAATGCTATAGTTACTCATTTAGATCATCTCC 781

QY 1174 AAAAATTTTATTAACCAATTTTATGTTTAAAAAATAA 1215
DB 782 AAAAATTTTATTAACCAATTTTATGTTTAAAAAATAA 823

RESULT 8
BX370934
LOCUS
DEFINITION BX370934 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS00C024YC05 5-PRIME, mRNA sequence.
ACCESSION BX370934
VERSION BX370934.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BH004ZB09QPI&cluster=7624.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BH004ZE09QPI.
FEATURES
source
1..948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00C024YC05"
/tissue="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 62.5%; Score 759.2; DB 13; Length 948;
Best Local Similarity 99.8%; Pred. No. 3.6e-164;
Matches 759; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 454 TCAGAGTACCTTCGGAGTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGC 513
DB 66 TCAGAGTACCTTCGGAGTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGC 125

QY 514 AGTTTTCGAGCTCCAGGAACCTTAAGAAATTAAGAACTTGACGATATGAATTCATTCAT 573
DB 126 AGTTTTCGAGCTCCAGGAACCTTAAGAAATTAAGAACTTGACGATATGAATTCATTCAT 185

QY 574 TTCTTCGAGGATAGACCACTTAATAAAGCTTCGCTTTTGTTCAGCAGCTCGAAATAAGC 633
DB 186 TTCTTCGAGGATAGACCACTTAATAAAGCTTCGCTTTTGTTCAGCAGCTCGAAATAAGC 245

QY 634 TTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAAATCTTGGATCTTTTGGAA 693
DB 246 TTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAAATCTTGGATCTTTTGGAA 305

QY 694 ATACTTTGAACCAACCAAGGCTCTCCAGTATAAAGCTGCAAGCACCATTAACTTTAT 753
DB 306 ATACTTTGAACCAACCAAGGCTCTCCAGTATAAAGCTGCAAGCACCATTAACTTTAT 365

QY 754 TGGAACTTTCTGACGAAACCAATATACATAATAGGATTCATATGCTCTCATATCATTC 813
DB 366 TGGAACTTTCTGACGAAACCAATATACATAATAGGATTCATATGCTCTCATATCATTC 425

QY 814 CATTCATCTCTGCAAGATTTGGATACCGCAAAATTTGTTTGTGCAAGATCTGTGC 873
DB 426 CATTCATCTCTGCAAGATTTGGATACCGCAAAATTTGTTTGTGCAAGATCTGTGC 485

QY 874 TGAACCTTTCAATCAAGAACTTACTCATGAAATCTGCATTTCTGTGCCACACCTGTGC 933
DB 486 TGAACCTTTCAATCAAGAACTTACTCATGAAATCTGTGCATTTCTGTGCCACACCTGTGC 545

QY 934 TCTTAGTAGAATATTTGGTGTGCTAGAGCACCATTATCTCTTATTTCTGTCTCTAG 993
DB 546 TCTTAGTAGAATATTTGGTGTGCTAGAGCACCATTATCTCTTATTTCTGTCTCTAG 605

QY 994 GCTGTATGTAAATTCCTCTGATATGTTAAAGTATGGGTGAGACCAAGAAAGAAATTT 1053
DB 606 GCTGTATGTAAATTCCTCTGATATGTTAAAGTATGGGTGAGACCAAGAAAGAAATTT 665

QY 1054 CAATAACAGATCAGTTTGGGTGCAATGATGATTTTTCAGCGCTCAAAATGGAGTAAGGA 1113

Db 666 CAATAACAGATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAAAATTGGAGTDAAGGA 725

QY 1114 AGATTCTGTATACCTCTCGAGAGGAGGATGTATAGTTACTTCATTTAGATGACTCC 1173

Db 726 AGATTCTGTATACCTCTCGAGAGGAGGATGTATAGTTACTTCATTTAGATGACTCC 785

QY 1174 AAAACCTTTTATAAACCCTTTTGTAGTTTAAAAAAGAAAA 1215

Db 786 AAAACCTTTTATAAACCCTTTTGTAGTTTAAAAAAGAAAA 827

RESULT 9
BU852149
LOCUS
DEFINITION AGENCOURT 10402016 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6617621
5', mRNA sequence.

ACCESSION BU852149.1 GI:24037112
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 791)
COMMENT NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaphs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LICM2863 row: d column: 05
High quality sequence stop: 588.

FEATURES
Location/Qualifiers
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6617621"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

ORIGIN
Query Match 62.4%; Score 758; DB 13; Length 791;
Best Local Similarity 99.2%; Pred. No. 6.9e-164;
Matches 761; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 449 TACACTCCAGAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCC 508

Db 3 TACACTCCAGAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCC 62

QY 509 TGTGAGTTTCCAGCTCCAGAACTTAAGATTTAAACTTTCAGATTAATGAT 568

Db 63 TGTGAGTTTCCAGCTCCAGAACTTAAGATTTAAACTTTCAGATTAATGAT 122

QY 569 TCAATTTCTTTCAGATAGGACAACTTAATAACCTTTGCTTTGTCAGCAGCTCGAA 628

Db 123 TCAATTTCTTTCAGATAGGACAACTTAATAACCTTTGCTTTGTCAGCAGCTCGAA 182

QY 629 TAAGCTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATACTTGGATCTTTT 688

Db 183 TAAGCTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATACTTGGATCTTTT 242

QY 689 TGGAAATACITTTTGAACAAACCAAAAGTCTTCCAGTAAATAAAGCTTGAAGCAACCAATTAAC 748

Db 243 TGGAAATACITTTTGAACAAACCAAAAGTCTTCCAGTAAATAAAGCTTGAAGCAACCAATTAAC 302

QY 749 TTTATTCGAATCTTCTGCAAGAACCATATTAACAATAAGGATTCATATGCTCTCATAT 808

Db 303 TTTATTCGAATCTTCTGCAAGAACCATATTAACAATAAGGATTCATATGCTCTCATAT 362

QY 809 CATTCACATTCATCTCTGCCAAGATTTGGATACCGCAAAATTTGTCTTCTGGAAGATT 868

Db 363 CATTCACATTCATCTCTGCCAAGATTTGGATACCGCAAAATTTGTCTTCTGGAAGATT 422

QY 869 CTGTCTGAACCTCTTTCATTCAGGAACCTACTACCATGAATCTGCATTCCTGTGGCCACAC 928

Db 423 CTGTCTGAACCTCTTTCATTCAGGAACCTACTACCATGAATCTGCATTCCTGTGGCCACAC 482

QY 929 TGTGCTCTTAGTAGTAATTTGGTGGTCTACTGAAGCACTTATATCTTATTTCTGTTC 988

Db 483 TGTGCTCTTAGTAGTAATTTGGTGGTCTACTGAAGCACTTATATCTTATTTCTGTTC 542

QY 989 TCTAGGCTGTATAGTCTAAATTTCTCTGATATCTTAAAGTAAATGGCTGAGACCAGAAAGA 1048

Db 543 TCTAGGCTGTATAGTCTAAATTTCTCTGATATCTTAAAGTAAATGGCTGAGACCAGAAAGA 602

QY 1049 AATTTCAATAACAGATCAGTTTGGGTGCATGTATGAATTTGACGCTCAAATTTGGAGTA 1108

Db 603 AATTTCAATAACAGATCAGTTTGGGTGCATGTATGAATTTGACGCTCAAATTTGGAGTA 662

QY 1109 AGGAGATTTCTGTACTTCTGCGAGAGAGCAATCTGTATAGTTACTCATTTAGATG 1168

Db 663 AGGAGATTTCTGTACTTCTGCGAGAGAGCAATCTGTATAGTTACTCATTTAGATG 722

QY 1169 ACTCCAAAACCTTTTATAAACCCTTTTGTAGTTTAAAAAAGAAAA 1215

Db 723 ACTCCAAAACCTTTTATAAACCCTTTTGTAGTTTAAAAAAGAAAA 769

RESULT 10
BX370935
LOCUS
DEFINITION BX370935 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC024YC05 5-PRIME, mRNA sequence.

ACCESSION BX370935
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 901)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAH0342E09RM1cluster=7624.f. Contact:
Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0BAH004ZE09RM1.

FEATURES
Location/Qualifiers
1..901
/organism="Homo sapiens"

ORIGIN		Query Match		Score		DB 13;		Length		901;	
		Best Local Similarity		97.68;		Pred. No. 2.8e-158;					
		Matches		742;		Conservative		0;		Mismatches	
				18;		Indels		0;		Gaps	
QY	454	TCGAGAGTCAC	TTGGAGTTTGGACCTCAGCAGAA	CAAAATCAAGGCAC	TCCCTGTGC	513					
DB	76	TCGAGAGTCAC	TTGGAGTTTGGACCTCAGCAGAA	CAAAATCAAGGCAC	TCCCTGTGC	135					
QY	514	AGTTTCCGAGCTCCAGGA	CTTAAGAAATTTAAACTTGGAGTAATGAATTCAT	CAAT	573						
DB	136	AGTTTCCGAGCTCCAGGA	CTTAAGAAATTTAAACTTGGAGTAATGAATTCAT	CAAT	195						
QY	574	TTCCCTTCGAAAGTAGGA	CAACTAATAAACCTTCGCTTTTGTCCAGCAGCTCGAAATAAGC	633							
DB	196	TTCCCTTCGAAAGTAGGA	CAACTAATAAACCTTCGCTTTTGTCCAGCAGCTCGAAATAAGC	255							
QY	634	TTCCATTTTGGCCTAGTGA	ATTTAGAAATTTTCCCTTTGAATAC	TTGGAA	693						
DB	256	TTCCATTTTGGCCTAGTGA	ATTTAGAAATTTTCCCTTTGAATAC	TTGGAA	315						
QY	694	ATACTTTTGAACCAAC	CAAAAGTCCTCCAGTAAATAAGCTGCAAGCACCATTAACTTTAT	753							
DB	316	ATACTTTTGAACCAAC	CAAAAGTCCTCCAGTAAATAAGCTGCAAGCACCATTAACTTTAT	375							
QY	754	TGGAATCTTCTGCAG	CAACCATATTAATAGGATTCATATGGCTCTCATATCATTC	813							
DB	376	TGGAATCTTCTGCAG	CAACCATATTAATAGGATTCATATGGCTCTCATATCATTC	435							
QY	814	CATTCCATCTCTCC	AGATTTCGATACCGCAAAATTCGTGTTGTTGGAGATCTGTC	873							
DB	436	CATTCCATCTCTCC	AGATTTCGATACCGCAAAATTCGTGTTGTTGGAGATCTGTC	495							
QY	874	TGAACCTCTTTCAT	TCAAGGAACCTACTACCAATGAATCTGATTCGTTGCCACACATCTGG	933							
DB	496	TGAACCTCTTTCAT	TCAAGGAACCTACTACCAATGAATCTGATTCGTTGCCACACATCTGG	555							
QY	934	TCTTAGTAGATAAT	TTGGGTGGTACTGAAGCCTATTATCTCTTATTTCTGTTCTCTCG	993							
DB	556	TCTTAGTAGATAAT	TTGGGTGGTACTGAAGCCTATTATCTCTTATTTCTGTTCTCTCG	615							
QY	994	GCTGTTATGTATA	TTTCCCTCTGATATGTAAAGTAAATGGGTGAGACCAAGAAAAAGAAATTT	1053							
DB	616	GCTGTTATGTATA	TTTCCCTCTGATATGTAAAGTAAATGGGTGAGACCAAGAAAAAGAAATTT	675							
QY	1054	CAATAACAGATCAG	TTGGGTGCATGTATGATTTTGCAGCGTCAATTTGCAGTAGAGGA	1113							
DB	676	GAATAACAGATCAG	TTGGGTGCATGTATGATTTTGCAGCGTCAATTTGCAGTAGAGGA	735							
QY	1114	AGATTTCTGTATAC	TTGCTGCAGAGGAGAAATGTGTATAGTTACTCATTTAGATGACTCC	1173							
DB	736	AGATTTCTGTATAC	TTGCTGCAGAGGAGAAATGTGTATAGTTACTCATTTAGATGACTCC	795							
QY	1174	AAAACTTTTATTA	AAAAACCAATTTTAGTTTAAAAA	1213							
DB	796	AAAACTTTTATTA	AAAAACCAATTTTAGTTTAAAAA	835							

EU621671
 EU621671.1 GI:23287886
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 739)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: James Martin
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@iowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 1-43, xAT rich#Low_complexity (matched complement)
 Seq primer: M.3 FORWARD
 POLYA=yes.
 FEATURES
 Location/Qualifiers
 1..739
 /organism="Homo sapiens"
 /molecule="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FLI-bga-b-24-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_FLI"
 /note="Organ: Chondrosarcoma; Vector: pT73-Pac
 (Pharmacia) with a modified polylinker; Site 1: Ecor I;
 Site 2: Not I; NCI CGAP FLI is a normalized cDNA library
 derived from a pool of mRNA obtained from 4 cell lines
 from grade III chondrosarcoma tissues. The library was
 constructed according to Bernaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 GAGGTCGGTG. The cell lines were provided by Dr. James
 Martin from the University of Iowa.
 TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LIB=UI-H-FLI

ORIGIN

	Query Match	58.4%;	Score 710;	DB 13;	Length 739;
	Best Local Similarity	99.6%;	Pred. No. 7.6e-153;		
	Matches 732;	Conservative	0;	Mismatches 1;	Indels 2; Gaps 2;
QY	481	TCAGCAAGAACAAAATCAAGGCACCTCCCTGTGCAGATTTTGCAGACTTTTGCAGAGCTCCAGGAACCTTAAGA	540		
DB	739	TCAGCAAGAACAAAATCAA-GCACTCCCTGTGCAG-TTTGCCAGCTNCAGGAACCTTAAGA	682		
QY	541	ATTTAAACCTTCGACGATATGATGATTCAATTCCTTGCAGATAGACACACTAATAA	600		
DB	681	ATTTAAACCTTCGCTATGATGATTCAATTCCTTGCAGATAGACACACTAATAA	622		
QY	601	ACCTTCGCTTTTGTGCAGCAGCTCGAAATAGACTTCATTTTGGCTAGTGAATTTAGAA	653		
DB	621	ACCTTCGCTTTTGTGCAGCAGCTCGAAATAGACTTCATTTTGGCTAGTGAATTTAGAA	562		
QY	661	ATTTATCCCTCGAATACTTGGACCTTTTGGAAATACTTTTGAAACCAACCAAAAGTCTTTC	720		


```

Db      126  GTTTGGGGTCAATGATGTTTGCAGCGTCATATTCGAGTAAGGAGAGATTTTGTATA 67
QY      1127  CTTGCTGGAGAGAGAAATGTTGATATAGTCTCAATTTAGTAGAGCTCCAAACATTTTATT 1185
Db      66   TTTGTTGGAGTGTGGTAATGTTGTTGTTGTTTATTATTATTTGTTGTTTATTATTTTNTT 8

RESULT 13
BU620411/c
LOCUS   BU620411
DEFINITION
  UI-R-FL1-bfw-1-09-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
  UI-R-FL1-bfw-1-09-0-UI 3', mRNA sequence.
ACCESSION
  BU620411
VERSION
  BU620411.1 GI:23286626
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 722)
  NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: James Martin
  cDNA library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@iowa.edu
  The following repetitive elements were found in this cDNA
  sequence: 1-43, SAT rich#Low_complexity (matched complement)
  Seq primer: M13 FORWARD
  POLVA=Yes.

Location/Qualifiers
  1..722
  /organism="Homo sapiens"
  /mol_type="RNA"
  /db_xref="taxon:9606"
  /clone="UI-R-FL1-bfw-1-09-0-UI"
  /tissue_type="Cell lines"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="NCI CGAP FL1"
  /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
  (pharmacia) with a modified polylinker; Site 1: EcoR I;
  Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
  derived from a pool of mRNA obtained from 4 cell lines
  from grade III chondrosarcoma tissues. The library was
  constructed according to Bonaldo, Lennon and Soares,
  Genome Research, 6:791-806, 1996. First strand cDNA
  synthesis was primed with an oligo-dT primer containing a
  Not I site. Double stranded cDNA was ligated to an EcoR I
  adaptor, digested with Not I, and cloned directionally
  into pT7T3-Pac vector. The oligonucleotide used to prime
  the synthesis of first-strand cDNA contains a library tag
  sequence that is located between the Not I site and the
  (drl18 tail. The sequence tag for this library is
  GAGGTCGGTG. The cell lines were provided by Dr. James
  Martin from the University of Iowa.
  TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
  TAG LIB=UI-R-FL1
  TAG_SEQ=GAGGTCGGTG"

FEATURES
  source

ORIGIN
  Query Match 57.9%; Score 763.8; DB 13; Length 722;
  Best Local Similarity 99.0%; Pred. No. 2e-151;
  Matches 768; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      501  GCATCCCTGTGCAGTTTGGCAGCTCCAGGAACCTTAAGAAATTTAAACTTGCAGCATAT 560

```

```

Db      721  GCATCCCTGTGCAGTTTGGCAGCTCCAGGAACCTTAAGAAATTTAAACTTGCAGCATAT 662
QY      561  GAATTAATCAATTTCTTGCAGAGTAGGCAACTAATAAACCTTCCTCTTTTGTACAGA 620
Db      661  GAATTAATCAATTTCTTGCAGAGTAGGCAACTAATAAACCTTCCTCTTTTGTACAGA 602
QY      621  GCTGGAATAAGCTTCCATTTTGCCTAGTAGTAATTTAGAAATTTATCCCTGGAATCTG 680
Db      601  GCTGGAATAAGCTTCCATTTTGCCTAGTAGTAATTTAGAAATTTATCCCTGGAATCTG 542
QY      681  GATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCTCTCCAGTAATAAAGCTCAAGA 740
Db      541  GATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCTCTCCAGTAATAAAGCTCAAGA 482
QY      741  CCAATTAATCTTTATGGAATCTTCTGCAACCAACCATATTACATAATAGGATCCCATGGC 800
Db      481  CCAATTAATCTTTATGGAATCTTCTGCAACCAACCATATTACATAATAGGATCCCATGGC 422
QY      801  TCTCATATCATCTCCATCTCCAGAGATTTGGATACCCGCAAAATTTGTGTTGT 860
Db      421  TCTCATATCATCTCCATCTCCAGAGATTTGGATACCCGCAAAATTTGTGTTGT 362
QY      861  GGAAGATTTCTGCTGAACTCTTTTCATTCAAGGAACCTACTACCATGAATCTGCACTCTG 920
Db      361  GGAAGATTTCTGCTGAACTCTTTTCATTCAAGGAACCTACTACCATGAATCTGCACTCTG 302
QY      921  GCCCACACTGTGCTTAGTAGATAATTTGGTGGTACTGGAACACTATTATCTCTAT 980
Db      301  GCCCACACTGTGCTTAGTAGATAATTTGGTGGTACTGGAACACTATTATCTCTAT 242
QY      981  TTTCTGTTCTTAGGCGTTATGTTTAACTCTCTGATATGTTAAAGTAATGGGTGAGACCA 1040
Db      241  TTTCTGTTCTTAGGCGTTATGTTTAACTCTCTGATATGTTAAAGTAATGGGTGAGACCA 182
QY      1041  GAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGTCATGATGATTTTCAGCGTCAAA 1100
Db      181  GAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGTCATGATGATTTTCAGCGTCAAA 122
QY      1101  TTGGAGTAAGGGAAGATTTCTGTATATCTGCTGGAGGAGGAATGTGTATAGTTACTCA 1160
Db      121  TTGGAGTAAGGGAAGATTTCTGTATATCTGCTGGAGGAGGAATGTGTATAGTTACTCA 62
QY      1161  TTTAGATGACTCCAAACCTTTTATTAAACCAATTTAGTTTAAAAAANAAAA 1215
Db      61   TTTAGATGACTCCAAACCTTTTATTAAACCAATTTAGTTTAAAAAANAAAA 7

BX367719      891 bp      mRNA      linear      EST 08-MAY-2003
BX367719 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS00B005YM21 5-PRIME, mRNA sequence.
BX367719
BX367719.1 GI:30437265
BX367719
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 891)
  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  Full-length cDNA libraries and normalization
  Unpublished (2001)
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  Library was constructed by Life Technologies, a division of
  Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
  Faraday Avenue Genoscope sequence ID : CS0AT005ZA07_T0433_1.
  Location/Qualifiers

```

FEATURES

```
source
1. .891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0B005YM21"
/issue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 57.4%; Score 700; DB 13; Length 891;
Best Local Similarity 97.0%; Pred. No. 1.5e-150;
Matches 733; Conservative 0; Mismatches 19; Indels 4; Gaps 2;

QY 460 AGTCACCTCGGAGTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAATTT 519
DB 20 AGTCACCTCGGAGTTGGACCT---CGCAAGACAAAATCAAGCACTCCCTGTGCAATTT 76
QY 520 GCACGCTCAGGAACCTAAGATTTAAATTTCAAGTATGATTAATGATTCATTTCCCTT 579
DB 77 GCACGCTCAGGAACCTAAGATTTAAATTTCAAGTATGATTAATGATTCATTTCCCTT 136
QY 580 GCAAGATAGCAACCTAATAAACCCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTCCAT 639
DB 137 GCAGATAGCAACCTAATAAACCCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTCCAT 196
QY 640 TTTTGCTAGTGAATTTAGAAATTTATCCCTTGAATTAATTTGGAATCTTTTGGAAATCTT 699
DB 197 TTTTGCTAGTGAATTTAGAAATTTATCCCTTGAATTAATTTGGAATCTTTTGGAAATCTT 256
QY 700 TTGAACAACCAAAAGCTCTCCAGTAAATTAAGCTGCAAGCACCATTAACTTTTGAAT 759
DB 257 TTGAACAACCAAAAGCTCTCCAGTAAATTAAGCTGCAAGCACCATTAACTTTTGAAT 316
QY 760 CTTCTGCAGCAACCATATTACATAATAGATTCATATGCTCTCATATCATTCATTC 819
DB 317 CTTCTGCAGCAACCATATTACATAATAGATTCATATGCTCTCATATCATTCATTC 376
QY 820 ATCTCTGCAAGATTTGGATACCGCAAAAATTTGCTTTTGTGGAAGATCTTGTCTGAAT 879
DB 377 ATCTCTGCAAGATTTGGATACCGCAAAAATTTGCTTTTGTGGAAGATCTTGTCTGAAT 436
QY 880 CTTTCAATCAAGGAACCTACTACCATGAATCTGATCTGTTGCGCCACACTGCTGTAG 939
DB 437 CTTTCAATCAAGGAACCTACTACCATGAATCTGATCTGTTGCGCCACACTGCTGTAG 496
QY 940 TAGATAAATTTGGTGTGCTAGTGAAGCACTATTATCTCTTATTTCTGTTCTAGGCTGTT 999
DB 497 TAGATAAATTTGGTGTGCTAGTGAAGCACTATTATCTCTTATTTCTGTTCTAGGCTGTT 556
QY 1000 ATGTTAATTTCTGATGTTTAAAGTAAATGCTGAGACAGAAAAGAAATTTCAATAA 1059
DB 557 ATGTTAATTTCTGATGTTTAAAGTAAATGCTGAGACAGAAAAGAAATTTCAATAA 616
QY 1060 CAGATCAGTTTGGGGTGCATGATGATTTTGGACGCTCAAAATGGAGTAAAGGAATTT 1119
DB 617 CAGATCAGTTTGGGGTGCATGATGATTTTGGACGCTCAAAATGGAGTAAAGGAATTT 676
QY 1120 CTGTATATCTGCTGGAGAGGAATGTTATAGTTACTCAATTTAGATGACTCCAAACT 1179
DB 677 CTGTATATCTGCTGGAGAGGA-CAATGTTGATAGTACTCAATTTAGATGACTCCAAACT 735
QY 1180 TTATTTAAACCAATTTTGTGTTTAAATAAAAAA 1215
DB 736 TTATTTAAACCAATTTTGTGTTTAAATAAAAAA 771

RESULT 15
BX394855/c
LOCUS BX394855 1165 bp mRNA linear EST 13-MAY-2003
```

```
DEFINITION BX394855 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC024YC05 3-PRIME, mRNA sequence.
ACCESSION BX394855
VERSION BX394855.1 GI:30628345
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1165)
Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC024AB03NP1&cluster=7624.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC024AB03NP1.
FEATURES
Location/Qualifiers
1..1165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YC05"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 57.3%; Score 695.8; DB 13; Length 1165;
Best Local Similarity 97.8%; Pred. No. 1.3e-149;
Matches 720; Conservative 6; Mismatches 10; Indels 2; Gaps 2;

QY 454 TCCAGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 513
DB 741 TCCAGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 682
QY 514 AGTTTGGCAGCTCCAGGAACCTTAAGATTTAAATCTTGACGATATGAATGATTCAT 573
DB 681 AGTTTGGCAGCTCCAGGAACCTTAAGATTTAAATCTTGACGATATGAATGATTCAT 622
QY 574 TCCCTTGCAGATAGGACCACTAATAAACCCTTCGCTTTTGTGACGAGCTCGAAATAGC 633
DB 621 TCCCTTGCAGATAGGACCACTAATAAACCCTTCGCTTTTGTGACGAGCTCGAAATAGC 562
QY 634 TCCCATTTTTCCTAGTGAATTTAGAAATTTATCCCTTGGAATCTTGGATCTTTTGGAA 693
DB 561 TCCCATTTTTCCTAGTGAATTTAGAAATTTATCCCTTGGAATCTTGGATCTTTTGGAA 502
QY 694 ATACTTTTGAACCAACCAAGTCTTCCAGTAAATAAGCTGCAAGCAACCAATTAACCTTAT 753
DB 501 ATACTTTTGAACCAACCAAGTCTTCCAGTAAATAAGCTGCAAGCAACCAATTAACCTTAT 442
QY 754 TGGAAATCTTCTGCAGCAACCAATATACATAAGGATTCATATGCTCTCATATCATTC 813
DB 441 TGGAAATCTTCTGCAGCAACCAATATACATAAGGATTCATATGCTCTCATATCATTC 382
QY 814 CATTTCCATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATTCGTGC 873
DB 381 CATTTCCATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATTCGTGC 322
QY 874 TGAATCTTTTCAATCAAGGAACCTACTACCAATGAACTGTGATTTCTGTTGCCCACTGTGG 933
```

```
Db 321 TGAACCTCTTCATTCAAGGAACACTACCATGAATCTGCATTCTGTTGCCACACTGTGG 262
Qy 934 TCCTAGTAGATAAATTGGGTGGTACTGAGCACCTATTATCTCTTATTCTCTGTTCTCTAG 993
Db 261 TCCTAGTAGATAAATTGGGTGGTACTGAGCACCTATTATCTCTTATTCTCTGTTCTCTAG 202
Qy 994 GCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGTGAGACCAGAAAAAGAAATTT 1053
Db 201 GCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGTGAGACCAGAAAAAGAAATTT 142
Qy 1054 CAATAACAGATCAGTTTGGGTGGCAATGATGATTTTGCAGGTCMAATGGAGTAAGGGA 1113
Db 141 CAATAACAGATCAGTTTGGGTGGCAATGATGATTTTGCAGGTCMAATGGAGTAAGGGA 82
Qy 1114 AGATTTCGTATACTTGTGGAGAGGGAATGTGTA-TAGTTACTC-ATTTAGTAGACT 1171
Db 81 AGATTTCGTATACTTGTGGAGAGGGAATGTGTANTAGTTACTCGATTTAGATKATT 22
Qy 1172 CCAAAACCTTTTATTAAAA 1189
Db 21 TTAANNNTWGHADDARD 4
```

Search completed: June 16, 2004, 00:57:25
Job time : 2403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 23:13:58 ; Search time 418 Seconds
(without alignments)
13267.648 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagcatgtatga.....ttagtttttaaaaaaaaaa 1215

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2998549 seqs, 2282253817 residues

Total number of hits satisfying chosen parameters: 5997098

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgm2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
3: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgm2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgm2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
19: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153.2	94.9	1471	15	US-10-037-270-1099
2	1153.2	94.9	1471	16	US-10-117-722-1099
3	950.2	78.2	1251	16	US-10-312-088-22
4	777	64.0	780	16	US-10-312-088-21
5	715.8	58.9	1241	9	US-09-833-381-1401
6	521.6	42.9	682	9	US-09-813-358-33
7	521.6	42.9	682	10	US-09-937-279-33
8	344	28.3	487	13	US-10-085-783A-7737
9	344	28.3	487	16	US-10-242-535A-7737
10	302.8	24.9	564	15	US-10-029-386-7642
11	281.4	23.2	732	13	US-10-076-555-463
12	281	23.1	732	13	US-10-076-555-467
13	239	19.7	239	15	US-10-029-386-21342
14	87.2	7.2	582	9	US-09-778-927A-28

15	74.8	6.2	172	13	US-10-085-783A-5829
16	74.8	6.2	172	16	US-10-242-535A-5829
17	60.2	5.0	2950	13	US-10-220-120-21
18	55	4.5	693	16	US-10-369-493-34864
19	55	4.5	4984	17	US-10-468-334-15
20	50	4.1	50	16	US-10-131-827-7581
21	49.8	4.1	2413	13	US-10-342-887-1742
22	49.8	4.1	2413	13	US-10-172-118-1742
23	48.2	4.0	1947	9	US-09-909-320-184
24	48.2	4.0	1947	9	US-09-909-088B-184
25	48.2	4.0	1947	9	US-09-905-291A-184
26	48.2	4.0	1947	9	US-09-902-853-184
27	48.2	4.0	1947	9	US-09-907-824-184
28	48.2	4.0	1947	9	US-09-907-841-184
29	48.2	4.0	1947	10	US-09-904-011-184
30	48.2	4.0	1947	10	US-09-906-742-184
31	48.2	4.0	1947	10	US-09-906-838-184
32	48.2	4.0	1947	10	US-09-907-613-184
33	48.2	4.0	1947	10	US-09-907-942-184
34	48.2	4.0	1947	10	US-09-904-859-184
35	48.2	4.0	1947	10	US-09-908-204-184
36	48.2	4.0	1947	10	US-09-904-820-184
37	48.2	4.0	1947	10	US-09-904-786-184
38	48.2	4.0	1947	10	US-09-906-646-184
39	48.2	4.0	1947	10	US-09-906-700-184
40	48.2	4.0	1947	10	US-09-903-786-184
41	48.2	4.0	1947	10	US-09-902-903-184
42	48.2	4.0	1947	10	US-09-903-749A-184
43	48.2	4.0	1947	10	US-09-904-119-184
44	48.2	4.0	1947	10	US-09-904-956-184
45	48.2	4.0	1947	10	US-09-902-736-184

ALIGNMENTS

RESULT 1
US-10-037-270-1099
Sequence 1099, Application US/10037270
Publication NO. US20030104529A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyang
APPLICANT: Chen, Rui-Hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: DT-FL_genes Version 1.0
SEQ ID NO 1099
LENGTH: 1471
TYPE: DNA
ORGANISM: Homo sapiens

Qy	1018	TGTTAAAGTAATCGGTGAGACCCAGAAAAAGAAATTTCAATAACAGATCAGTTTGGGGGTGC	1071
Db	1274	TGTTAAAGTAATGGGTGAGACCCAGAAAAAGAAATTTCAATAACAGATCAGTTTGGGGGTGC	1333
Qy	1078	ATGTATGATTTTGCACCGTCAAATTCGAGTAAGGGGAAGATTTCTGTATACTTGTCTGGAGA	1137
Db	1334	ATGTATGATTTTGCACCGTCAAATTCGAGTAAGGGGAAGATTTCTGTATACTTGTCTGGAGA	1393
Qy	1138	GGAGGAATGTATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTTAAAAACCAATTTT	1197
Db	1394	GGAGGAATGTATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTTAAAAACCAATTTT	1453
Qy	1198	AGTTTTTAAAAAARAAAA 1215	
Db	1454	AGTTTTTAAAAAARAAAA 1471	
RESULT 2			
US-10-117-722-1099			
Sequence 1099, Application US/10117722			
Publication No. US20030219744A1			
GENERAL INFORMATION:			
APPLICANT: Tang, Y. Tom			
APPLICANT: Liu, Chenghua			
APPLICANT: Asundi, Vinod			
APPLICANT: Zhang, Jie			
APPLICANT: Dmanac, Radoje T.			
TITLE OF INVENTION: NO. US20030219744A1 Nucleic Acids and			
TITLE OF INVENTION: Polypeptides			
FILE REFERENCE: 784CIP2BCIP			
CURRENT APPLICATION NUMBER: US/10/117,722			
CURRENT FILING DATE: 2002-04-04			
PRIOR APPLICATION NUMBER: 09/620,312			
PRIOR FILING DATE: 2000-07-19			
PRIOR APPLICATION NUMBER: 09/552,317			
PRIOR FILING DATE: 2000-04-25			
PRIOR APPLICATION NUMBER: 09/468,725			
PRIOR FILING DATE: 2000-01-21			
NUMBER OF SEQ ID NOS: 1104			
SOFTWARE: pc PL_genes Version 1.0			
SEQ ID NO 1399			
LENGTH: 1471			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (40)..(1284)			
US-10-117-722-1099			
Query Match 94.9%; Score 1153.2; DB 16; Length 1471;			
Best Local Similarity 99.7%; Pred. No. 2e-286; 3; Indels 0; Gaps 0;			
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	58	TTATTCCAGGCCATTTCCAGCAGTTTAAAAAGGTTTCCTTTTCAGCTATGAGACTGGCTCATTA	117
Db	314	TAACTAAGGCCATTTCCAGCAGTTTAAAAAGGTTTCCTTTTCAGCTATGAGACTGGCTCATTA	373
Qy	118	GAGGCTGTAAATGTTGATACACAGTTTCAACCGCTACACCGTAGAGACTTCAGAAATTG	177
Db	374	GAGGCTGTAAATGTTGATACACCGTTTCAACCGCTACACCGTAGAGACTTCAGAAATTG	433
Qy	178	AAAACTTTAAACTAAATAGGTTATCACATCCAAAAGAGACTATCCTCTTAAGTAAGAAAT	237
Db	434	AAAACTTTAAACTAAATAGGTTATCACATCCAAAAGAGACTATCCTCTTAAGTAAGAAAT	493
Qy	238	TTCCATATTCCTTTGGAACATCTTCAGACTTCCTTACCTGTGGGCTTGTCCGAGTTGATATGC	297
Db	494	TTCCATATTCCTTTGGAACATCTTCAGACTTCCTTACCTGTGGGCTTGTCCGAGTTGATATGC	553
Qy	298	GTATGCTTTGCTTAAAAAGCCCTTAGGAAATTAGACTTGAGTCACAAACCATATAAAAAAGC	357
Db	554	GTATGCTTTGCTTAAAAAGCCCTTAGGAAATTAGACTTGAGTCACAAACCATATAAAAAAGC	613


```
QY 358 TTCCAGCTCAATTGAGAGCTCATACACCTTCAAGAACTTACCTGAATGCAATCACT 417
Db 614 TTCCAGCTCAATTGAGAGCTCATACACCTTCAAGAACTTACCTGAATGCAATCACT 673
QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCACTTACACTCCAGAGTCACTTCGGAGTTGG 477
Db 674 TGGAGTCATTTAGTGTAGCTTGTGTCACTTACACTCCAGAGTCACTTCGGAGTTGG 733
QY 478 ACTCAGCAAGAACAAAATCAAGGCACTCCTGTGTGAGTTTTCGAGCTCCAGGAACCTTA 537
Db 734 ACTCAGCAAGAACAAAATCAAGGCACTCCTGTGTGAGTTTTCGAGCTCCAGGAACCTTA 793
QY 538 AGAATTTAAACCTTGACGATAATGAATGATTCATTTCTCTGCAAGATAGGCAACTTA 597
Db 794 AGAATTTAAACCTTGACGATAATGAATGATTCATTTCTCTGCAAGATAGGCAACTTA 853
QY 598 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 657
Db 854 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 913
QY 658 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 717
Db 914 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 973
QY 718 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTTGGAATCTTCTGCAAGAACCATAT 777
Db 974 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTTGGAATCTTCTGCAAGAACCATAT 1033
QY 778 TACATAATAGGAATCCATATGCTCTCATATCATTCATTCATTCCTCCAGAGTTGG 837
Db 1034 TACATAATAGGAATCCATATGCTCTCATATCATTCATTCCTCCAGAGTTGG 1093
QY 838 ATACCGCAAAATTTGTCTTTGCGAGATCTCTGTGAACTCTTTCATTCAGGAACTA 897
Db 1094 ATACCGCAAAATTTGTCTTTGCGAGATCTCTGTGAACTCTTTCATTCAGGAACTA 1153
QY 898 CTACCATGAATCTGCATCTGTGTGCCACACCTGTGCTTGTAGTAAATTTGGGTGGTA 957
Db 1154 CTACCATGAATCTGCATCTGTGTGCCACACCTGTGCTTGTAGTAAATTTGGGTGGTA 1213
QY 958 CTGAAGCACTTATATCTTATTTCTGTCTCTAGGCTGTATGTATATCTCTGTGATA 1017
Db 1214 CTGAAGCACTTATATCTTATTTCTGTCTCTAGGCTGTATGTATATCTCTGTGATA 1273
QY 1018 TGTAAAGTAATGGTGTAGAGACAGAAAAGAAATTTCAATAACAGATCAAGTTGGGGTGC 1077
Db 1274 TGTAAAGTAATGGTGTAGAGACAGAAAAGAAATTTCAATAACAGATCAAGTTGGGGTGC 1333
QY 1078 ATGTATGATTTTGCAGCTGCAAAATTTGGAGTAAGGGAAGATTTCTGTATATCTTGTGAG 1137
Db 1334 ATGTATGATTTTGCAGCTGCAAAATTTGGAGTAAGGGAAGATTTCTGTATATCTTGTGAG 1393
QY 1138 GGAGGATGTGTATGATTTACTCATTTAGATGACTCCAAATCTTTTATTAACCAATTTT 1197
Db 1394 GGAGGATGTGTATGATTTACTCATTTAGATGACTCCAAATCTTTTATTAACCAATTTT 1453
QY 1198 AGTTTAAAAAATAAAAA 1215
Db 1454 AGTTTAAAAAATAAAAA 1471
```

RESULT 3

```
US-10-312-088-22
; Sequence 22, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabnic, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Smith, Randall F.
```

```
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Xie, Qing
; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-312-088-22
```

Query Match

Best Local Similarity 99.1%; Score 950.2; DB 16; Length 1251;

Matches 968; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```
QY 58 TTATTCAGGCCATTTCCAGCGATTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 117
Db 275 TAAGTAAGGCCATTTCCAGCGATTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 334
QY 118 GAGGCTGTAATGTGTATACACAGTTTCAAGCTCACACAGTGAAGACTTCAGAAATTG 177
Db 335 GAGGCTGTAATGTGTATACACAGTTTCAAGCTCACACAGTGAAGACTTCAGAAATTG 394
QY 178 AAAACCTTTAAAACTAAATGGTTATCCACATCCAAAAAGACTATCCTCAAGTAAGATT 237
Db 395 AAAACCTTTAAACTAAATGGTTATCCACATCCAAAAAGACTATCCTCAAGTAAGATT 454
QY 238 TTCCATATTTCTTGGAACTCTTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATA 297
Db 455 TTCCATATTTCTTGGAACTCTTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATA 514
QY 298 GTATCGCTTTGTCTTAAAAAGCCTTAGGAAATTAGACTTAGACTCAACACATATAAAAAAGC 357
Db 515 GTATCGCTTTGTCTTAAAAAGCCTTAGGAAATTAGACTTAGACTCAACACATATAAAAAAGC 574
QY 358 TTCCAGCTCAATTTGGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAACTCACT 417
Db 575 TTCCAGCTCAATTTGGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAACTCACT 634
QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCACTTCTACACTCCAGAGTCACTTCGGAGTTGG 477
Db 635 TGGAGTCATTTAGTGTAGCTTGTGTCACTTCTACACTCCAGAGTCACTTCGGAGTTGG 694
QY 478 ACTCAGCAAGAACAAAATCAAGGCACTCCTGTGTGAGTTTTCGAGCTCCAGGAACCTTA 537
Db 695 ACTCAGCAAGAACAAAATCAAGGCACTCCTGTGTGAGTTTTCGAGCTCCAGGAACCTTA 754
QY 538 AGAATTTAAACCTTGACGATAATGAATGATTCATTTCTCTGCAAGATAGGCAACTTA 597
Db 755 AGAATTTAAACCTTGACGATAATGAATGATTCATTTCTCTGCAAGATAGGCAACTTA 814
QY 598 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 657
Db 815 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 874
QY 658 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 717
Db 875 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 934
QY 718 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTTGGAATCTTCTGCAAGAACCATAT 777
Db 935 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTTGGAATCTTCTGCAAGAACCATAT 994
```



```

Db 491 AAAACCTTAAACCTAAATGCTTATCATCCAAAAGAGCATCTCTTAAGTAAGATT 550
QY 238 TTCCATATTCCTTTGGAAACATCTTCAGACTTCTTACTGTGGGCTGTGCGAGTTGATATGC 297
Db 551 TTCCATATTCCTTTGGAAACATCTTCAGACTTCTTACTGTGGGCTGTGCGAGTTGATATGC 610
QY 298 GTATGCTTTGCTTTAAAGAGCTTAGGAAATAGACTTGTAGCTCAACCATATATAAAGC 357
Db 611 GTAIGCTTTGCTTTAAAGAGCTTAGGAAATAGACTTGTAGCTCAACCATATATAAAGC 670
QY 358 TTCCAGCTCAAAATGGAGAGCTTCATACRCCTTCAAGAACTTAACTGGAATSAACAATCACT 417
Db 671 TTCCAGCTCAAAATGGAGAGCTTCATACRCCTTCAAGAACTTAACTGGAATSAACAATCACT 730
QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTCTACCTCCAGAAAGTCACTTCGGAGTTGG 477
Db 731 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTCTACCTCCAGAAAGTCACTTCGGAGTTGG 790
QY 478 ACCTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCAAGTTTTCGAGCTCCAGGAACCTTA 537
Db 791 ACCTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCAAGTTTTCGAGCTCCAGGAACCTTA 850
QY 538 AGAATTTAAACACTTGA-CGATAATGAATTAATCAATTTCCCTTCAAGATAGGACAACTA 596
Db 851 AGAATTTAAACACTTGAACCGATATGAATTAATCAATTTCCCTTCAAGATAGGACAACTA 910
QY 597 ATAAACCTTCGCTTTTGTGTCAGGAGCTCGAAATAGCTTCCATTTTTCCTAGTGAATTT 656
Db 911 ATAAACCTTCGCTTTTGTGTCAGGAGCTCGAAATAGCTTCCATTTTTCCTAGTGAATTT 970
QY 657 AGAATTTATCCCTTGAATCTTGAATCTTTTGGAAATTAATTTGAACCAACCAAAAGTC 716
Db 971 AGAATTTATCCCTTGAATCTTGAATCTTTTGGAAATTAATTTGAACCAACCAAAAGTC 1030
QY 717 CTTCAGTAAATAAGCTCAAGCAACCAATTAATTTATTTGAATCTTCTGCAAGCAACATA 776
Db 1031 CTTCAGTAAATAAGCTCAAGCAACCAATTAATTTATTTGAATCTTCTGCAAGCAACATA 1090
QY 777 TTACATAATAGGATTCATATGCTCTCTCATATCAT 811
Db 1091 TTACATAATAGGATTAAGATTTTAAATAGTCAATGTAAT 1125

```

```

RESULT 6
US-09-813-358-33
; Sequence 33, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813,358
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-358-33

```

```

Query Match 42.9%; Score 521.6; DB 9; Length 682;
Best Local Similarity 97.5%; Pred. No. 7.2e-124;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

```

```

QY 277 GSCCTGTCCGAGTTGATATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 336
Db 1 GSCCTGTCCGAGTTGATATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 60
QY 337 GTCAACCAACATATAAAAAGCTTCCAGCTACAATTTGGAGAGCTCATACACCTTCAAGAAC 396
Db 61 GTCAACCAACATATAAAAAGCTTCCAGCTACAATTTGGAGAGCTCATACACCTTCAAGAAC 420
QY 397 TTAACCTGAATGCAATCACTTGGAGTCATTTAGTGTAGCTTGTGTCAATTCACATCC 456
Db 121 TTAACCTGAATGCAATCACTTGGAGTCATTTAGTGTAGCTTGTGTCAATTCACATCC 480
QY 457 AGAAGTCACCTTCGAGTTTGGAGCTCAGCAAGAACCAAAATCAAGGCACTCCCTGTGAGT 516
Db 181 AGAAGTCACCTTCGAGTTTGGAGCTCAGCAAGAACCAAAATCAAGGCACTCCCTGTGAGT 540
QY 517 TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAACCTTGAAGATTAATGAATTCATTTTC 576
Db 241 TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAACCTTGAAGATTAATGAATTCATTTTC 600
QY 577 CTTGCAAGATAGGACAACTAATAAACCCTTGCCTTTTGTGAGGAGCTCGAAATAGCTTTC 636
Db 301 CTTGCAAGATAGGACAACTAATAAACCCTTGCCTTTTGTGAGGAGCTCGAAATAGCTTTC 660
QY 637 CATTTTGGCTTGTGAAATTTAGAAATTTATCCCTTGAATTAATTTGAAATTAATTTGG 696
Db 361 CATTTTGGCTTGTGAAATTTAGAAATTTATCCCTTGAATTAATTTGAAATTAATTTGG 720
QY 697 CTTTGAACCAACCAAAAGCTCCTTCCAGTAATAAAGCTGCAAGCAACCAATTAATTTTC 756
Db 421 CTTTGAACCAACCAAAAGCTCCTTCCAGTAATAAAGCTGCAAGCAACCAATTAATTTTC 780
QY 757 AATCTTCTGCAAGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 812
Db 481 AATCTTCTGCAAGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 840
QY 813 CCATTCATCTCTGCCAAGATTGG 837
Db 541 CCATTCATCTCTGCCAAGATTGG 865

```

```

RESULT 7
US-09-997-279-33
; Sequence 33, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997,279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-997-279-33

```

```

Query Match 42.9%; Score 521.6; DB 10; Length 682;
Best Local Similarity 97.5%; Pred. No. 7.2e-124;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

```

```

QY 277 GSCCTGTCCGAGTTGATATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 336
Db 1 GSCCTGTCCGAGTTGATATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 60

```

337	Qy	GTCAACACCATATATAAAGAGCTTCAGCTACAAATGGAGAGCTCATACACCTTCAAGAAC	396
61	Db	GTCAACACCATATATAAAGAGCTTCAGCTACAAATGGAGAGCTCATACACCTTCAAGAAC	120
397	Qy	TTAACCTGGAATGACAATCACATCTGGAGTCAATTTAGTGTAGCCCTTGTGTCACTACCTCC	456
121	Db	TTAACCTGGAATGACAATCACATCTGGAGTCAATTTAGTGTAGCCCTTGTGTCACTACCTCC	180
457	Qy	AGAAAGTCACTTCGGAGTTTGGACCTTCAGCAAGACAAAAATCAAGGCATCTCCCTGFGCAGT	516
181	Db	AGAAAGTCACTTCGGAGTTTGGACCTTCAGCAAGACAAAAATCAAGGCACCTCCCTGFGCAGT	240
517	Qy	TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAAACTTGACGATAATGAATTCGAATTC	576
241	Db	TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAAACTTGACGATAATGAATTCGAATTC	300
577	Qy	CTTGCAAGATAGGCACTAATAAACCTTCGCTCTTTGTACGAGCTCGAAATAAGCTTC	636
301	Db	CTTGCAAGATAGGCACTAATAAACCTTCGCTCTTTGTACGAGCTCGAAATAAGCTTC	360
637	Qy	CATTTTGCCTTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATA	696
361	Db	CATTTTGCCTTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATA	420
697	Qy	CTTTTGAACCAACCAAAAAGTCTCTCGAGTAATAAAGCTGCAGAGCACCAATTAATTTGG	756
421	Db	CTTTTGAACCAACCAAAAAGTCTCTCGAGTAATAAAGCTGCAGAGCACCAATTAATTTGG	480
757	Qy	AATCTTCTGCAGCAACCATATTTACATAAT--AGGATTCATATGGCTCT--CATATCAT	812
481	Db	AATCTTCTGCAGCAACCATATTTACATAATAGGATTCATATGGCTCTCTCATATTCATTT	540
813	Qy	CCATTCATCTCTGCCAAGATTGG	837
541	Db	CCATTCATCTCTGCCAGNATTGG	565

RESULT 5

```

US-10-085-783A-7737
; Sequence 7737, Application US/10065783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7737
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-7737

```

QY	846	AAAATTGTGTTTGTGGAGAGATTCTGTCGAACTCTTTTCATTCCAGGAACTACTACCATG	905
Db	204	AAAATTGTGTTTGTGGAGAGATTCTGTCGAACTCTTTTCATTCCAGGAACTACTACCATG	263
QY	906	AAATCTGCATTCTGTCGCCACACACTGTGGCTTTAGTAGATAAATTTGGGTGCTACTGAAGCA	965
Db	264	AAATCTGCATTCTGTCGCCACACACTGTGGCTTTAGTAGATAAATTTGGGTGCTACTGAAGCA	323
QY	966	CCATTATATCTCTTATTTCTGTTCTCTAGGCTGTTATGTGTTAAATTCCTCTGATATGTTAAAG	1025
Db	324	CCATTATATCTCTTATTTCTGTTCTCTAGGCTGTTATGTGTTAAATTCCTCTGATATGTTAAAG	383
QY	1026	TAAATGGGTGAGACGACGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCAATGATGA	1085
Db	384	TAAATGGGTGAGACGACGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCAATGATGA	443
QY	1086	TTTTGCAGCGTCAAAATTGGAGTGAAGGAGATTTCTGTATACCTT	1129
Db	444	TTTTGCAGCGTCAAAATTGGAGTGAAGGAGATTTCTGTATACCTT	487

RESULT 9
 US-10-242-535A-7737
 ; Sequence 7737, Application US/10242535A
 ; Publication No. US20040013663A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ChondroGene Inc.
 ; APPLICANT: Liem, C.C.
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 ; FILE REFERENCE: 4231/2005
 ; CURRENT APPLICATION NUMBER: US/10/242,535A
 ; CURRENT FILING DATE: 2002-09-12
 ; PRIOR APPLICATION NUMBER: US 10/085,783
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/305,340
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/275,017
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: US 60/271,955
 ; PRIOR FILING DATE: 2001-02-28
 ; NUMBER OF SEQ ID NOS: 58994
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7737
 ; LENGTH: 487
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (42)..(42)
 ; OTHER INFORMATION: n is a, c, g, or t
 US-10-242-535A-7737

RESULT 9

```

US-10-242-535A-7737
; Sequence 7737, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/0895,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,343
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7737
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-7737

```

	Query Match	28.3%	Score 344	DB 13	Length 487
	Best Local Similarity	100.0%			
	Matches 344	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	786	AGGATTCATATGGCTCTCATATCATTCATTCATCTGCCAAGATTGGATACGCA	845		
Db	144	AGGATTCATATGGCTCTCATATCATTCATTCATCTGCCAAGATTGGATACGCA	203		

RESULT 11
US-10-076-555-463
; Sequence 463, Application US/10076555
; Publication No. US20030065156A1
; GENERAL INFORMATION:

Query Match	23.2%	Score 281.4	DB 13	Length 750
Best Local Similarity	87.0%	Pred. No. 6e-62		
Matches 320	Conservative 0	Mismatches 45	Indels 3	Gaps 2
Qy	786	AGGATTCGATATGGCTCTCATATCATTCATTCCTCTCTGCCAAGATTTCGGHATACGGCA	845	
Db	383	AGGATTCGATATGGCTCTCATATCATTCATTCCTCTCTGCCAAGATTTCGGHATACGGCA	442	
Qy	846	AAATTTGTGTTGTGGAGATTCTCTCGAACTCTTTTCATTCGAAGAACTACTACCATG	905	
Db	443	AAATTTGTGTTGNGGAGATTCTCNCCTGAACTCTTTCATTCGAAGAACTACTACCATG	502	
Qy	906	AATCTGCATCTGTGTCCACACTGTGGCTTAGTAGATAATTTGGGTGGTACTGAAGCA	965	
Db	503	AATCTGCATCTGTGTCCACACTGTGGCTTAGTAGATAATTTGGGTGGTACTGAAGCA	560	
Qy	966	CTTATTATCTCTATTCTTGTTCTCTPAGCTGTGTAATTCCTCTGATATGTT-AAA	1024	
Db	561	CTTATTATCTCTATTCTTGTTCTCTPAGCTGTGTAATTCCTCTGATATGTT-AAA	620	
Qy	1025	GTAATGGGTGAGACACAGAAAAAGAAATTCATATACAGATCAGTTTGGGGTGCATGTATG	1084	
Db	621	GTAATGGGTGAGACACAGAAAAAGAAATTCATATACAGATCAGTTTGGGGTGCATGTATG	680	

RESULT 14

US-09-778-927A-28
; Sequence 28, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(582)
; OTHER INFORMATION: n = a,c,g,t any unknown or other
US-09-778-927A-28

Query Match 7.2%; Score 87.2; DB 9; Length 582;
Best Local Similarity 83.1%; Pred. No. 5.7e-12;
Matches 123; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

Qy	28	CAGTGGATTATATTTACAACTTTATTTATTTATTCAGGCCATTTCCAGCAGTTTAAAG	87
Db	436	GAATGAAATGCTACAGCTATACCCAGACCTTTTATAGGCCATTTCCAGCAGTTTAAAG	495
Qy	88	GTTT-CTTTTCAGCTATGAGTCTCATAGAGGCTGTAAATGTTTCATACACAGTTTCA	146
Db	496	GTTTCCCTTTTCAGCTATGAGTCTCATAGAGGCTGTAAATGTTTCATACACAGTTTCA	554
Qy	147	ACGTCACACAGTGAAGACTTCAGAAAT	174
Db	555	ACGTCACACAGTGAAGACTTCGAAT	582

RESULT 15

US-10-085-783A-5829
; Sequence 5829, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5829
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-5829

Query Match 6.2%; Score 74.8; DB 13; Length 172;
Best Local Similarity 74.7%; Pred. No. 4.5e-09;
Matches 124; Conservative 0; Mismatches 32; Indels 10; Gaps 2;

Qy	1060	CAGATCAGTTTGGGCTGCAATGATATGATTTTGCAGCGTCAAAATTTGAGTAGGGAAGATTT	1119
Db	1	CAGATCAGTTTGGGCTGCAATGATATGATTTTGCAGCGTCAAAATTTGAGTAGGGAAGATTT	60
Qy	1120	CTGTATA-----CTTGCTGGAGAGGAGGAATGTGTATAGTTTACTCATTTA-----CATGA	1169
Db	61	CTTGCTATATCTTTGCTCGGGAGAGGAGGAATGTGTATAGAACTGATTTAGGGTGACTT	120
Qy	1170	CTCCAAAACCTTTTATATAAACCAATTTAGTTTAAARAAAAA	1215
Db	121	CCAAAACCTTTTGTATAAACCAATTTAGTTTTCAGAAAGATAA	166

Search completed: June 16, 2004, 01:06:24
Job time : 421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 23:00:52 ; Search time 81 seconds
(without alignments)
8324.263 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagcatgtatgta.....ttagttttaaaaaaaaaa 1215

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCRUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153.2	94.9	1471	4	US-09-620-312D-1099
2	715.8	58.9	1241	4	US-09-833-381-1401
3	56.2	4.6	2169	4	US-09-434-408-3
4	48.2	4.0	1947	4	US-09-907-794A-184
5	48.2	4.0	1947	4	US-09-905-125A-184
6	48.2	4.0	1947	4	US-09-902-775A-184
7	44.2	3.6	832	4	US-09-621-976-2813
8	44	3.6	1749	4	US-09-081-149-5
9	43.2	3.6	7218	1	US-08-232-463-14
10	42	3.5	419	4	US-09-833-381-1373
11	41.2	3.4	1881	4	US-09-833-381-1406
12	39.4	3.2	1680	4	US-09-081-149-1
13	38.4	3.2	832	4	US-09-621-976-2813
14	38.4	3.2	3401	4	US-09-907-794A-249
15	38.4	3.2	3401	4	US-09-905-125A-249
16	38.4	3.2	3401	4	US-09-902-775A-249
17	38.4	3.2	4428	4	US-09-833-381-1364
18	38.2	3.1	450	4	US-09-621-976-406
19	38.2	3.1	495	4	US-09-312-283C-361
20	38.2	3.1	5199	4	US-09-620-312D-40
21	37.8	3.1	7304	4	US-10-234-708-44
22	37.6	3.1	9370	1	US-08-322-559-27
23	37.6	3.1	9370	3	US-08-545-860D-27
24	37.6	3.1	9370	5	PCT-US94-04496-27
25	37.6	3.1	9391	1	US-08-320-559-25
26	37.6	3.1	9391	3	US-08-545-860D-25
27	37.6	3.1	9391	5	PCT-US94-04496-25

C 28	37.4	3.1	3026	4	US-09-963-137-136	Sequence 136, App
C 29	37.4	3.1	3026	4	US-09-963-137-140	Sequence 140, App
C 30	37	3.0	1749	4	US-09-081-149-6	Sequence 6, Appli
C 31	36.2	3.0	1218	4	US-09-252-991A-14561	Sequence 14561, A
C 32	36.2	3.0	5407	3	US-09-269-040-7	Sequence 7, Appli
C 33	36.2	3.0	6877	1	US-08-347-340-1	Sequence 1, Appli
C 34	36	3.0	72604	4	US-09-268-992-7	Sequence 7, Appli
C 35	36	3.0	72604	4	US-09-657-474-7	Sequence 7, Appli
C 36	36	3.0	640681	4	US-09-790-988-1	Sequence 1, Appli
C 37	35.8	2.9	260	4	US-09-833-381-557	Sequence 557, App
C 38	35.8	2.9	842	4	US-09-484-370B-16	Sequence 16, Appl
C 39	35.6	2.9	3012	4	US-09-184-418C-57	Sequence 57, Appl
C 40	35.6	2.9	8700	2	US-08-392-625-16	Sequence 16, Appl
C 41	35.6	2.9	8750	2	US-08-466-961A-16	Sequence 16, Appl
C 42	35.6	2.9	8700	2	US-08-645-193B-18	Sequence 18, Appl
C 43	35.6	2.9	8954	4	US-09-184-418C-6	Sequence 6, Appli
C 44	35.6	2.9	640681	4	US-09-790-988-1	Sequence 1, Appli
C 45	35.4	2.9	11459	4	US-09-462-136-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-1099
; Sequence 1099, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 1399
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1284)
US-09-620-312D-1099

Query Match 94.9%; Score 1153.2; DB 4; Length 1471;
Best Local Similarity 99.7%; Pred. NO. 1.2e-312;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTATTCAGGCCATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 117
Db 314 TAAGTAAGGCCATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 373

```
QY 118 GAGCGTGAATGTTGATACACCGAGTTTCAACGCTCACACCGAGTGAAGACTTTCAGAAATTG 177
Db 374 GAGCGTGAATGTTGATACACCGAGTTTCAACGCTCACACCGAGTGAAGACTTTCAGAAATTG 433
QY 178 AAAACCTTTAAACCTAAATGGTTATCATCATCCAAAAAGACTATCCTCTAAGTAAGAAATT 237
Db 434 AAAACCTTTAAACCTAAATGGTTATCATCATCCAAAAAGACTATCCTCTAAGTAAGAAATT 493
QY 238 TTCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTCCGAGTTGATATGC 297
Db 494 TTCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTCCGAGTTGATATGC 553
QY 298 GTATGCTTGTCTTTAAAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAAAGC 357
Db 554 GTATGCTTGTCTTTAAAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAAAGC 613
QY 358 TTCAGCTACAAATGGAGACTCATACACCTTCAAGAACTTAACTGTAAGTCAATCACT 417
Db 614 TTCAGCTACAAATGGAGACTCATACACCTTCAAGAACTTAACTGTAAGTCAATCACT 673
QY 418 TGGAGTCAATTTAGTGTAGCCTTGTCTCATCTCACTCCAGAGTCACTTGGAGTTGG 477
Db 674 TGGAGTCAATTTAGTGTAGCCTTGTCTCATCTCACTCCAGAGTCACTTGGAGTTGG 733
QY 478 ACCTCAGCAAGAACAAATCAAGGCACCTCCCTGTGCAAGTTTGTCCAGTCCAGGAACCTTA 537
Db 734 ACCTCAGCAAGAACAAATCAAGGCACCTCCCTGTGCAAGTTTGTCCAGTCCAGGAACCTTA 793
QY 538 AGAATTTAAACCTTGAGTAATGAATGAATTCATTTCCCTGCAAGATAGGACAACTAA 597
Db 794 AGAATTTAAACCTTGAGTAATGAATGAATTCATTTCCCTGCAAGATAGGACAACTAA 853
QY 598 TAAACCTTCGCTTTTGTGACGAGCTCGAATAGCTTCCATTTTCCCTAGTGAATTTA 657
Db 854 TAAACCTTCGCTTTTGTGACGAGCTCGAATAGCTTCCATTTTCCCTAGTGAATTTA 913
QY 658 GAAATTTATCCCTTGAATATCTGATCTTTTGGAAATATCTTTGAAACCAACCAAGTCC 717
Db 914 GAAATTTATCCCTTGAATATCTGATCTTTTGGAAATATCTTTGAAACCAACCAAGTCC 973
QY 718 TTCCAGTAATAAGCTGCAAGCAGCAATTAATTTATTTGGAATCTTCTGCAACCAACCAAT 777
Db 974 TTCCAGTAATAAGCTGCAAGCAGCAATTAATTTATTTGGAATCTTCTGCAACCAACCAAT 1033
QY 778 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATCTCTGCAAGATTTGG 837
Db 1034 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATCTCTGCAAGATTTGG 1093
QY 838 ATACCGCAAAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTCATTCAGAGAACTA 897
Db 1094 ATACCGCAAAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTCATTCAGAGAACTA 1153
QY 898 CTACCATGATCTGCAATCTGTGTCGCCACACTGCTGCTTAGTAGATTAATTTGGTGGTA 957
Db 1154 CTACCATGATCTGCAATCTGTGTCGCCACACTGCTGCTTAGTAGATTAATTTGGTGGTA 1213
QY 958 CTGAGCACTATATCTCTTATTTCTGTTCTCTAGCTGTGATGTTAATTTCTCTGATA 1017
Db 1214 CTGAGCACTATATCTCTTATTTCTGTTCTCTAGCTGTGATGTTAATTTCTCTGATA 1273
QY 1018 TGTAAAGTAATGGGTGAGACCAAGAAAAAGAAATTTCAATAACAGATCACTTTGGGGTGC 1077
Db 1274 TGTAAAGTAATGGGTGAGACCAAGAAAAAGAAATTTCAATAACAGATCACTTTGGGGTGC 1333
QY 1078 ATGTATGATTTTGCAGGCTCAATTTGGAGTAAAGGAAATTTCTGTATCTTCTGAGTA 1137
Db 1334 ATGTATGATTTTGCAGGCTCAATTTGGAGTAAAGGAAATTTCTGTATCTTCTGAGTA 1393
QY 1138 GGAGGAATGTATAGTTACTCATTTAGATGATCCAAAACCTTTTATTAACCAATTTT 1197
Db 1394 GGAGGAATGTATAGTTACTCATTTAGATGATCCAAAACCTTTTATTAACCAATTTT 1453
QY 1198 AGTTTTAAAAAATAAAAA 1215
```

```
Db 1454 AGTTTTAAAAAATAAAAA 1471
RESULT 2
US-09-833-381-1401
; Sequence 1401, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1401
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1401
Query Match 58.9%; Score 715.8; DB 4; Length 1241;
Best Local Similarity 97.6%; Pred. No. 2e-190;
Matches 737; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
QY 58 TTATTCAGGCGCAATTCAGCAGATTTAAAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 117
Db 371 TAAGTAAGGCGCAATTCAGCAGATTTAAAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 430
QY 118 GAGCTGTAATGTTGATACACCGAGTTTCAACGCTCACACCGAGTGAAGACTTTCAGAAATTG 177
Db 431 GAGCTGTAATGTTGATACACCGAGTTTCAACGCTCACACCGAGTGAAGACTTTCAGAAATTG 490
QY 178 AAAACCTTTAAACCTAAATGGTTATCATCATCCAAAAAGACTATCCTCTAAGTAAGAAATT 237
Db 491 AAAACCTTTAAACCTAAATGGTTATCATCATCCAAAAAGACTATCCTCTAAGTAAGAAATT 550
QY 238 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGCTTGTCCGAGTTGATATGC 297
Db 551 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGCTTGTCCGAGTTGATATGC 610
QY 298 GTATGCTTGTCTTTAAAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAAAGC 357
Db 611 GTATGCTTGTCTTTAAAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAAAGC 670
QY 358 TTCCAGCTACAAATGGAGACTCATACACCTTCAAGAACTTAAACCTGAAATGACAACTACT 417
Db 671 TTCCAGCTACAAATGGAGACTCATACACCTTCAAGAACTTAAACCTGAAATGACAACTACT 730
QY 418 TGAAGTCAATTTAGTGTAGCCTTGTCTCATCTACCTCCAGAAAGTCACTTCGAGTTTGG 477
Db 731 TGAAGTCAATTTAGTGTAGCCTTGTCTCATCTACCTCCAGAAAGTCACTTCGAGTTTGG 790
QY 478 ACCTCAGCAGACCAATTCAGGCACTCCCTGTGCAAGTTTGTCCAGTCCAGGAACCTTA 537
Db 791 ACCTCAGCAGACCAATTCAGGCACTCCCTGTGCAAGTTTGTCCAGTCCAGGAACCTTA 850
QY 538 AGAATTTAAACCTTGA- CGAATATGAATTTGATTAATTTCTTGCAGATAGGACAACTA 596
Db 851 AGAATTTAAACCTTGA- CGAATATGAATTTGATTAATTTCTTGCAGATAGGACAACTA 910
QY 597 ATAAACCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTTCATTTTGTGCTAGTGAATTT 656
Db 911 ATAAACCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTTCATTTTGTGCTAGTGAATTT 970
QY 657 AGAATTTATCCCTTCAATACTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTGC 716
Db 971 AGAATTTATCCCTTCAATACTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTGC 1030
QY 717 CTTCAGTAATAAAGCTGCAAGCACCATTAACTTTATTTGNAATCTTCTTCGACGAACCTA 776
```

Db 1031 CTTCCAGTAATAAGCTGCAAGCACCATTAATTTATTGGAATCTTCTGCAAGAACCATTA 1090
QY 777 TTACATATAGATTCATATATGCTTCATATCAT 811
Db 1091 TTACATAATAGTAAGATTTTAAATAGTCATGTAAT 1125
RESULT 3
US-09-434-408-3
; Sequence 3, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPO3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of zapop3
; NAME/KEY: misc.feature
; LOCATION: (1)...(2169)
; OTHER INFORMATION: n = A,T,C or G
US-09-434-408-3
Query Match 4.6%; Score 56.2; DB 4; Length 2169;
Best Local Similarity 30.4%; Pred. No. 6.5e-06;
Matches 109; Conservative 57; Mismatches 183; Indels 9; Gaps 1;
QY 301 TCGTTGCTTAAAGCCCTTAGGAATAGACTTGAGTCACACACCATATATAAAGCTTC 360
Db 230 TTTTWSNTYNGCNACNATHAAGTNTYNGAYTNCAYGAYTNCAYTNCNGCNTNC 289
QY 361 CAGCTACATATGAGACCTCATACACCTTCAAGACTTAACCTSAATGACATCACTTGG 420
Db 290 CNGAYGYTNGCNCARYTNCNGCNTNCARGTNTYNAAYGNGARMGNAAYCARYTNA 349
QY 421 AGTCATTTAGTGTAGCTTGTGTCATCTACCTCAGAGTCACCTGGAGTTGGACC 480
Db 350 TGCARYTNCNMGNSNATHGGNAAYTNCNCA-----RYTNCARACNTYNAAYG 400
QY 481 TCAGCAAGAACAAATCAAGGCACTCCCTGTGCGAGTTTTCAGCTCCAGGAACCTTAAGA 540
Db 401 TAAAGAYAAAYATYNAARGARYTNCNGAYCNGCTNGNGARYTNGWSNTYNGNA 460
QY 541 ATTTAAACTTCACGATATATGATGATCAATTTCTTCAGATAGGACACTAATAA 600
Db 461 CNYTNAAYATHWSNGGNAAYGARATHCARMGNTNCNCARATGYTNGCNCAYTNGNA 520
QY 601 ACCTTCGCTTTTGTGACAGCTCGAAATAGCTTCATTTTGCCTAGTGAATTAG 658
Db 521 CNYTNGARATGYTNSYNTNGAYGNCNCGNATGTTNTAYCCNCGMNGARGTNTG 578

RESULT 4
US-09-907-794A-184
; Sequence 184, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Inc

; APPLICANT: Eator, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Pilvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30993
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 184
; LENGTH: 1947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;
QY 248 CTTGGAACTCTTCACACTTCTACTGTGGCTGTCCGAGTTCGATATCGTATGCTTTG 307

Db 1121 CTGGAGTCACTTATTTCTCTAACCAAGCTCGAATCTTACAGTGGCAGTATTAG 1180
Qy 308 CTTAAAGCCTTAGGAATAGACTGAGTCAACCAATATAAAGAGCTTCCAGCTAC 367
Db 1181 TTTACAGAACTCAGTGTCTAGATGTGAGCTCAACCAATTTCAATGATTTCAATAGA 1240
Qy 368 AATTGGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTTGGAGTCAAT 427
Db 1241 AATAGGATGCTTTCAGAACCTTGAGCAATTTGCATATCACTGGGAACAAGTGGACATTTCT 1300
Qy 428 TAGTGTAGCCTTGTGTGATCTTCACTCCAGAACTCACTTGGAGTGGACCTCAGCAA 487
Db 1301 GCCAAACAATTTGTTAAATGCATAAAGTTGA-----GGACTTTTGAATCTGGGACA 1351
Qy 488 GAACAAATCAAGCACTCCCTGTGCGAGTTTGGCAGCTCCAGGAACCTTAAGAATTTAAA 547
Db 1352 GAATGATCACTCACTCCAGAAAGTTGGTCACTCTCCAGCTCACTCAGCTGGA 1411
Qy 548 ACTTGACGATAAATGAATGATTCAATTTCCCTTGCAGATAGGACA 592
Db 1412 GCTGAAGGGGAACTGCTTGGACCGCTGCCAGCCAGCTGGGCA 1456

RESULT 5

US-09-905-125A-184
; Sequence 184, Application US/09905125A

; Patent No. 6664376

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 184
; LENGTH: 1947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-905-125A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

Qy 248 CTTGGAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGCGTATGCTTTG 307
Db 1121 CTTGAGTCACTTATTTCTCTAACCAAGCTCGAATCTTACAGTGGCAGTATTAG 1180
Qy 308 CTTAAAGCCTTAGGAATAGACTGAGTCAACCAATATAAAGAGCTTCCAGCTAC 367
Db 1181 TTTACAGAACTCAGTGTCTAGATGTGAGCTCAACCAATTTCAATGATTTCAATAGA 1240
Qy 368 AATTGGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTTGGAGTCAAT 427
Db 1241 AATAGGATGCTTTCAGAACCTTGAGCAATTTGCATATCACTGGGAACAAGTGGACATTTCT 1300
Qy 428 TAGTGTAGCCTTGTGTGATCTTCACTCCAGAACTCACTTGGAGTGGACCTCAGCAA 487
Db 1301 GCCAAACAATTTGTTAAATGCATAAAGTTGA-----GGACTTTTGAATCTGGGACA 1351
Qy 488 GAACAAATCAAGCACTCCCTGTGCGAGTTTGGCAGCTCCAGGAACCTTAAGAATTTAAA 547
Db 1352 GAATGATCACTCACTCCAGAAAGTTGGTCACTCTCCAGCTCACTCAGCTGGA 1411
Qy 548 ACTTGACGATAAATGAATGATTCAATTTCCCTTGCAGATAGGACA 592
Db 1412 GCTGAAGGGGAACTGCTTGGACCGCTGCCAGCCAGCTGGGCA 1456

RESULT 6

US-09-902-775A-184

; Sequence 184, Application US/09902775A

; Patent No. 6686451

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

; APPLICANT: Goddard, A.

APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Pami, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 184
LENGTH: 1947
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

QY 248 CTTGAACATCTTCAGACTCTTACTGTGGGCTTGCCGAGTTGATGCGTATGCTTTG 307
DB 1121 CTTGAGTCATCTTATTTCTCTAACACAGCTCGAATCTTACAGTGGCAGTATTTAG 1180

QY 308 CTTAAAAAGCCTTAGGAATTAGACTTCAGTCAACACCAATATAAAAAAGCTTCCAGCTAC 367
DB 1181 TTTACAGAACTCAGATGCTTAGATGTGAGCTACACCAACATTTCAATGATTTCCATAGA 1240

QY 368 AATGGAGACCTCATAACCTTCAAGAACTTAACTGATGCAATCACTCTGGAGTCAAT 427
DB 1241 AATAGGATTTGCTTCAGAACCTGCAGCATTTGCAATATCACTGGGAACAAAGTGCATCT 1300

QY 428 TAGTGAGCCTTGCTGCTCATTCTACACTCCAGAGTCACTTCGAGTTTGGACCTAGCAA 487
DB 1301 GCCAAACAATTTGTTAAATGCATAAAGTTGA-----GGACTTTGAATCTGGGACA 1351
QY 488 GAACAAAATCAAGGCACTCCCTGTGCAGTTTGGCCAGTCTCCAGGAACCTTAAGAATTTAA 547
DB 1352 GAATGCACTCACTCACTCCAGAGAAAGTTGGTCACTCTCCAGCTCACTCAGCTGGA 1411
QY 548 ACTGACGATATGAATTAATTCATTCCTTGGCAAGATAGACA 592
DB 1412 GCTGAAGGGGAACCTGCTGGACCGCTGCCAGCCAGCTGGGCCA 1456

RESULT 7

US-09-621-976-2813
Sequence 2813, Application: US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

Query Match

3.6%; Score 44.2; DB 4; Length 832;

Best Local Similarity 10.8%; Pred. No. 0.0094;
Matches 39; Conservative 176; Mismatches 139; Indels 7; Gaps 1;

QY 854 TGTGTGGAAGATTCTGCTGAATCTTTCAATCAAGGAACCTACTACCAATGATCTGCA 913
DB 17 TKKWSWSYMWYKWKYMKTYWNRKXKKKAWKYKWTWYWRYYMYWYKWKYKWKAMCET 76
QY 914 TTCTGTTGCCACACTGTGGTCTTAGTAGATAATTTGGGTGCTACTGAAGCACCTATTAT 973
DB 77 KTKKKKKGYMMWYWGWRWNRWYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKY 136
QY 974 CTCCTATTCTCTCTAGGCTGTATCTGTAATCTCTGATATCTTAAAGTAATGSGT 1033
DB 137 SSKGWTWKKKKKANTTWWKKYWAATRYWMMWCKWKRASWYCWMMGKRWSTW 196
QY 1034 GAGACGAGAAAGAAATTTCAATCAAGATCAGTTGGGTGGTCAATGATTTTGGAG 1093
DB 197 RKRSYASARSARCC-----YSCWAGMSWKYMWNRWNRWATGAGMKAWRASCWM 249
QY 1094 COTCAATTTGAGTAAGGAGGAGATTCTGTATCTTCTGAGAGAGGAGGAATGTGTATAG 1153
DB 250 RKYAGSKTSYKSWMWNRWNRWYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKY 309
QY 1154 TPTACTATTAGATGACTCCAAACCTTTTATTAACCAATTTTAGTTTAAAAAATAA 1213
DB 310 WWCWARMYRSTGTGRASWNRWYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKYKWKY 369
QY 1214 A 1214
DB 370 A 370

RESULT 8

US-09-081-149-5
Sequence 5, Application US/0908149A

```
/ Patent No. 6506889
/ GENERAL INFORMATION:
/ APPLICANT: Sieburch, Derek
/ TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
/ FILE REFERENCE: UTC-02938
/ CURRENT APPLICATION NUMBER: US/09/081.149A
/ CURRENT FILING DATE: 1998-05-19
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patent in Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 1749
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ OTHER INFORMATION: The N at location 30 can be A, C, T, or G.
/ OTHER INFORMATION: The N at location 90 can be A, C, T, or G.
/ OTHER INFORMATION: The N at location 90 can be A, C, T, or G.
/ US-09-081-149-5

Query Match      3.6%; Score 44; DB 4; Length 1749;
Best Local Similarity 50.0%; Pred. No. 0.015;
Matches 144; Conservative 0; Mismatches 135; Indels 9; Gaps 1;

QY 317 CTTAGGAATTAAGCTTGAGTGCACACCATATAAAAAAGCTTCCAGCTACAAATTGGAGA 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 CTTCTACACCTTCAGAACCTTACCTGAATGATGATCACTTGGAGTCATTTAGTGTAGC 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 437 CTTGTGTCTTCTACACTCCAGAGTCACTTGGAGTTTGGACCTCAGCAAGCAAAAT 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 497 CRAAGCACTCCCTGTGCGAGTTTGGCAGTCCAGGACCTTAAGAAATTTAAACCTTGACGA 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 557 TAATGAATGATTCATTTCTTCCAGATAGGACACTTAATAACCT 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 559 GAACCTGCTCACTTCCCTCCAGGAATTCGGTACACTGGAAAAACCT 1556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-08-232-463-14
/ Sequence 14, Application US/08232463
/ Patent No. 5670367
/ GENERAL INFORMATION:
/ APPLICANT: DORNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313

/ Patent No. 6506889
/ GENERAL INFORMATION:
/ APPLICANT: Sieburch, Derek
/ TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
/ FILE REFERENCE: UTC-02938
/ CURRENT APPLICATION NUMBER: US/09/081.149A
/ CURRENT FILING DATE: 1998-05-19
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patent in Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 1749
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ OTHER INFORMATION: The N at location 30 can be A, C, T, or G.
/ OTHER INFORMATION: The N at location 90 can be A, C, T, or G.
/ OTHER INFORMATION: The N at location 90 can be A, C, T, or G.
/ US-09-081-149-5

Query Match      3.6%; Score 43.2; DB 1; Length 7218;
Best Local Similarity 8.1%; Pred. No. 0.049;
Matches 36; Conservative 21; Mismatches 198; Indels 0; Gaps 0;

QY 601 ACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCATTTTTCGCTAGTGAATTAGAA 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1036 AGCTTGCTCGAGTCCAGGAGCTTCGATATTTTTTGGAAATCTTTTGAACAACCAAGTCCCTC 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 ATTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGAACAACCAAGTCCCTC 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1096 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1155
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 CAGTAATTAAGCTCGAAGCACCATTAACTTTATTTGGAATCTCTGCAAGCAACCAATATAC 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1156 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1215
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 ATAATAGATTCCATATGGCTCTCATATCATTCATTCATCTCGCAAGATTGGATA 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1216 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1275
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 CCACAAATTTGTTTGTGGAGATCTCTCTGAACTCTTTTCATTCAGGAATCTACTA 900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1276 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1335
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 901 CCATGAATCTGATTCGTTGCCACACTCTGGTCTTAGTAGATAATTTGGTGGTACTG 960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1336 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1395
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 AAGCACCTATTATCTCTTATTTCTGTCTTAGCTGTATTTATTTCTCTCTGATATGT 1020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1396 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1455
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 TAAAGTAATGGGTGAGACACAGAAA 1044
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1456 TTAAGTACTTGATAGATAGTAA 1479
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-833-381-1373
/ Sequence 1373, Application US/09833381
/ Patent No. 6672186
/ GENERAL INFORMATION:
/ APPLICANT: Robison, Keith E.
/ TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
/ FILE REFERENCE: 5800-119
/ CURRENT APPLICATION NUMBER: US/09/833,381
/ CURRENT FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 09/516,448
/ PRIOR FILING DATE: 2000-02-29
/ NUMBER OF SEQ ID NOS: 2050
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1373
```

Query Match 3.5%; Score 42; DB 4; Length 419;
Best Local Similarity 60.5%; Pred. No. 0.028;
Matches 69; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

US-09-833-381-1373

314 AGCCCTTAGAATTAGACTTGAGTGCACACCTTAACCTGAATGCAATCACTTGGAGTCATT 427
374 AGCCCTTAGAATTAGACTTGAGTGCACACCTTAACCTGAATGCAATCACTTGGAGTCATT 427
306 AGCCCTTAGAATTAGACTTGAGTGCACACCTTAACCTGAATGCAATCACTTGGAGTCATT 427

RESULT 11

US-09-833-381-1406
Sequence 1406, Application US/09833381
Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith B.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1406
LENGTH: 1881
TYPE: DNA
ORGANISM: Homo sapiens

US-09-833-381-1406

Query Match 3.4%; Score 41.2; DB 4; Length 1881;
Best Local Similarity 53.8%; Pred. No. 0.095;
Matches 85; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

US-09-833-381-1406

305 TTGCTTAAAGCCCTTAGAATTAGACTTGAGTGCACACCTTAACCTGAATGCAATCACTTGGAGTC 424
339 TGGCCTCACACACCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 458
365 TACAACTGAGACCTCATACACCTTCAAGAACTTAACCTGAATGCAATCACTTGGAGTC 424
459 AGACTTGGCGCTGCTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 518
425 ATTTAGTGTAGCTTGT 462
519 CTGCTGT 556

RESULT 12

US-09-081-149-1
Sequence 1, Application US/09081149A
Patent No. 6506889
GENERAL INFORMATION:
APPLICANT: Han, Min
TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
FILE REFERENCE: UTC-02938
CURRENT APPLICATION NUMBER: US/09/081,149A
CURRENT FILING DATE: 1998-05-19
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 1680
TYPE: DNA
ORGANISM: Caenorhabditis elegans

US-09-081-149-1

Query Match 3.2%; Score 39.4; DB 4; Length 1680;
Best Local Similarity 65.2%; Pred. No. 0.29;
Matches 58; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

US-09-621-976-2813/c

340 ACAACCATATAAAAGCTTCCAGCTACAATTTGGAGACCTCATACACCTTCAAGACCTTA 399
1220 ACAATCACTGAAAAGCTTCCAAATCAATAGGAATCTCATATAAACTCCGGAGCTGG 1279
400 ACTGGAATCAATCACTTGGAGTCATT 428
1280 ATCTGAGGAATGAATGGAGACCTT 1328

RESULT 13

US-09-621-976-2813/c
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399

US-09-621-976-2813

Query Match 3.2%; Score 38.4; DB 4; Length 832;
Best Local Similarity 10.4%; Pred. No. 0.39;
Matches 36; Conservative 157; Mismatches 153; Indels 0; Gaps 0;

US-09-621-976-2813

850 TTGTGTTTGTGAAGATCTGTCTGAATCTTCTTCAATCAAGAACTTACCATGAATC 909
347 TWTMMWKKARRYYWKKSTVACASRYKRYTWGWWYWKWKMSTRWYCYWKKCCMY 288
910 TGCATCTCTTGGCCACACTGTGCTTGTAGTATTAATTTGGGTGGTGTCTGAAGACCTTA 969
287 RGRCAWYTWARGRWWSYAWGKWSKMSYMSCTRYKKSTYWTNKTCTCATWYWK 238
970 TTATCTCTTATTTCTGTCTCTAGGCTGTATGTTATTTCTCTGATATGTTAAAGTAAT 1029
227 YKRWWSKTCWSGSGYNTSYTSYSYWSYWSYWSYWSYWSYWSYWSYWSYWSYWSYWSY 165
1030 GGTGAGACCAAGAAATTTCAATAACAGATCAGTTTGGGTGTCATGTATGATTTT 1089
167 RYATTWRRAMWMAAATWMMYMAWCKSSRGAAMYRTTMMWGYRWKKSYYRTTC 108
1090 GCAGCGTCAATTTGGAGTGAAGGAAGATTTCTGTATATCTCTGGAGAGAGGAATGTGT 1149
107 AWAYAKTKRSYWCWKKKCKMMWMAAYGKTMWRACTWKTWYRWMAWMAWMAWMAW 48
1150 ATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTAACCAAT 1195
47 TMMWYRWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 2

RESULT 14

US-09-907-794A-249
Sequence 249, Application US/09907794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi


```
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 249
; LENGTH: 3401
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-249

Query Match      3.2%; Score 38.4; DB 4; Length 3401;
Best Local Similarity 53.3%; Pred. No. 0.75;
Matches 81; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY      453 CTCAGAAGTCACCTCGAGCTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTG 512
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1403 CTCCTGCAGAACCTCCAGAACCTAGCCATCAGGCCAACCGGATCGAGCGCTCCCTCG 1462
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      513 CAGTTTGGCAGCTCCAGGAACCTTAAGAAATTTAAACTTCACGATATGAATTGATCAA 572
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1463 GAGCTCTTCAGTGCCGGAAGCTGCGGCCCTGCACCTGGGCAACACGCTGCTGCAGTCA 1522
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      573 TTTCCTGCAAGATAGACAACTAATAAACCT 604
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1523 CTGCCCTCCAGGGTGGCGGAGCTGACCAACCT 1554
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 16, 2004, 00:59:05
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 19:53:27 ; Search time 399 Seconds
(without alignments)
12936.242 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagatgatgtgta.....ttagtttaaaaaaaaaaaaa 1215

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: geneseqm:980s: *
2: geneseqm:990s: *
3: geneseqm:2000s: *
4: geneseqm:2001as: *
5: geneseqm:2001bs: *
6: geneseqm:2002s: *
7: geneseqm:2003as: *
8: geneseqm:2003bs: *
9: geneseqm:2003cs: *
10: geneseqm:2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1215	100.0	1215	4	AAC66898 Human EXM
2	1153.2	94.9	1471	4	Aai59205 Human pol
3	1153.2	94.9	1471	8	Adb49189 Novel hum
4	1153	94.9	1600	6	Abk90911 CDNA enco
5	1139	93.7	1341	4	Aas33220 DNA enco
6	951.4	78.3	1287	4	Aai60991 Human pol
7	950.2	78.2	1251	6	Aad27818 Human leu
8	777	64.0	780	6	Aad27817 Human leu
9	521.6	42.9	682	4	Aas56409 Human CDN
10	443.8	36.5	467	6	Abi68766 Kidney ca
11	425.2	35.0	1204	7	Abz36181 Human ca
12	418.8	34.5	584	4	Aah09329 Human CDN
13	418.8	34.5	1590	4	Aai51533 Human CDN
14	331	27.2	333	7	Abz19104 Group III
15	281.4	23.2	750	2	Aax98737 Human val
16	281	23.1	732	2	Aax98741 Human val
17	188.8	15.5	770	2	Aai6653 Human gen
18	91.6	7.5	430	5	Aaf64752 Novel hum
19	87.2	7.2	582	6	Abz51842 Novel hum
20	85.2	7.0	382	5	Aaf64860 Novel hum
21	60.2	5.0	2950	4	Aas31006 Human dia
22	59.6	4.9	2156	7	Acd13337 Human DNA
23	56.2	4.6	2169	3	Aaa30203 Human RIN

ALIGNMENTS

RESULT 1	
AAC66898	
ID AAC66898 standard; cDNA; 1215 BP.	
XX	
AC AAC66898;	
XX	
DT 27-MAR-2001 (first entry)	
XX	
DE Human EXMAD-9 coding sequence SEQ ID NO: 34.	
XX	
XX Extracellular matrix and adhesion-associated protein; EXMAD; cancer;	
XX Inflammation; reproductive disorder; cardiovascular disorder;	
XX Immune disorder; musculoskeletal disorder; developmental disorder;	
XX Gastrointestinal disorder; cell proliferation disorder; ss.	
XX Homo sapiens.	
XX	
PN WO200068380-A2.	
XX	
PD 16-NOV-2000.	
XX	
PF 10-MAY-2000; 2000WO-US012811.	
XX	
PR 11-MAY-1999; 99US-0133643P.	
PR 23-AUG-1999; 99US-0150409P.	
XX	
PA (INCY-) INCYTE GENOMICS INC.	
XX	
PI Bandman O, Hillman JL, Tang YT, Lal P, Yue H, Baughn MR, Lu DM;	
PI Azimzai Y;	
XX	
XX WPI; 2001-007395/01.	
DR P-PSDB; AAB27231.	
XX	
XX Isolated polynucleotide encoding extracellular matrix or adhesion-	
PT associated protein (EXMAD) useful for diagnosing, treating, or preventing	
PT disorders associated with expression of EXMAD such as proliferative,	
PT immune and genetic disorders.	
PS Claim 4; Page 118; 129pp; English.	
XX	
XX The present invention provides the protein and coding sequences for 25	
CC novel extracellular matrix and adhesion-associated proteins (EXMADs).	
CC These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5, EXMAD-	
6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12, EXMAD-13,	
EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19, EXMAD-20,	

24	56.2	4.6	2169	6	ABS54163	Abz54163 Human deg
25	55	4.5	4984	6	AAL49657	Aal49657 Human neu
26	54.4	4.5	2694	3	AAL45671	Aaz45671 Nucleotid
27	53.4	4.4	1867	6	ABQ99593	Abq99593 Human cod
28	52.4	4.4	2620	4	AAL14529	Aal14529 Human CDK
29	50.8	4.2	2000	7	ADA71938	Ada71938 Rice gene
30	50	4.1	50	6	ABZ07590	Abz07590 Human leu
31	49.8	4.1	2410	4	AAL14179	Aal14179 Human CDN
32	49.8	4.1	2410	6	AAL49327	Aal49327 Human sig
33	48.2	4.0	1593	4	AAS33111	Aas33111 DNA enco
34	48.2	4.0	1630	4	AAL161075	Aal161075 Human pol
35	48.2	4.0	1947	2	AAX52247	Aax52247 Protein P
36	48.2	4.0	1947	3	ADC78504	Adc78504 Human PRO
37	48.2	4.0	1947	4	AAP72405	Aap72405 Human PRO
38	48.2	4.0	1947	7	ACA59047	Aca59047 Human PRO
39	48.2	4.0	1947	7	ACA58444	Aca58444 CDNA enco
40	48.2	4.0	1947	7	ACA60151	Aca60151 Human CDN
41	48.2	4.0	1947	7	ACD07551	Acd07551 Novel hum
42	48.2	4.0	1947	7	ABX71599	Abx71599 Human CDN
43	48.2	4.0	1947	7	ACH06931	Ach06931 Human sec
44	48.2	4.0	1947	7	ABX96188	Abx96188 Human sec
45	48.2	4.0	1947	7	ACA05489	Aca05489 CDNA enco

CC EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are useful in
 CC the prevention and treatment of cancers, cell proliferation,
 CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
 CC gastrointestinal disorders and inflammation
 XX
 SQ Sequence 1215 BP; 381 A; 231 C; 210 G; 393 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1215; DB 4; Length 1215;
 Best Local Similarity 100.0%; Pred. No. 2.3e-306;
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAAGAACTAGCATGTATATATCTCCAGTGGAAATTAATTAATCTCAACCTTTATTTA 60
 DB 1 GAAGAACTAGCATGTATATATCTCCAGTGGAAATTAATTAATCTCAACCTTTATTTA 60
 QY 61 TTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGCTGCTCATAGAG 120
 DB 61 TTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGCTGCTCATAGAG 120
 QY 121 GCTGTAATGTGTATACACCAAGTTTCAACGCTCCACCAAGTGAAGACTTCAGAAATTCGAA 180
 DB 121 GCTGTAATGTGTATACACCAAGTTTCAACGCTCCACCAAGTGAAGACTTCAGAAATTCGAA 180
 QY 181 ACTTTAAACTAAATGGTTTATCACATCCAAAAGAACTATCCTCTAAGTAAGAAATTTTC 240
 DB 181 ACTTTAAACTAAATGGTTTATCACATCCAAAAGAACTATCCTCTAAGTAAGAAATTTTC 240
 QY 241 CATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGCCGAGTTGATGCGTA 300
 DB 241 CATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGCCGAGTTGATGCGTA 300
 QY 301 TGCCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGAGTCAACCACTATATAAAGCTTC 360
 DB 301 TGCCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGAGTCAACCACTATATAAAGCTTC 360
 QY 361 CAGCTACAATTTGGAGACCTCATACACCTTCAAGAACTTAACTCTGAATGACAACTACCTGG 420
 DB 361 CAGCTACAATTTGGAGACCTCATACACCTTCAAGAACTTAACTCTGAATGACAACTACCTGG 420
 QY 421 AGTCATTTAGTGTAGCTTGTGTCATCTTCACTACCTCAGAAAGTCACTTCGGAGTTGGACC 480
 DB 421 AGTCATTTAGTGTAGCTTGTGTCATCTTCACTACCTCAGAAAGTCACTTCGGAGTTGGACC 480
 QY 481 TCAGCAAGCAACAAATCAAGGCACTCCCTGTGAGTTTCCAGCTCCAGAACTTAAGA 540
 DB 481 TCAGCAAGCAACAAATCAAGGCACTCCCTGTGAGTTTCCAGCTCCAGAACTTAAGA 540
 QY 541 ATTATAAACTTGACGATATGAATGATTCATTTCTTGCAGAGTAGGACAACTAATAA 600
 DB 541 ATTATAAACTTGACGATATGAATGATTCATTTCTTGCAGAGTAGGACAACTAATAA 600
 QY 601 ACCTTCGCTTTTGTGTCAGAGCTCGAAATGAAGTTCCTATTTTGCCTAGTGAATTTAGAA 660
 DB 601 ACCTTCGCTTTTGTGTCAGAGCTCGAAATGAAGTTCCTATTTTGCCTAGTGAATTTAGAA 660
 QY 661 ATTATPCCCTTGAATACCTTGATCTTTTGGAAATATCTTTGAAACCAACCAAAAGTCCCTC 720
 DB 661 ATTATPCCCTTGAATACCTTGATCTTTTGGAAATATCTTTGAAACCAACCAAAAGTCCCTC 720
 QY 721 CAGTAATAAGCTGCAAGCAACCAATTAATTTATGAACTTCTGCAAGCAACCAATTTAC 780
 DB 721 CAGTAATAAGCTGCAAGCAACCAATTAATTTATGAACTTCTGCAAGCAACCAATTTAC 780
 QY 781 ATAATAGGATTCATATGCTCTCATATCATCTCCATCTCTCCAGATTTGATA 840
 DB 781 ATAATAGGATTCATATGCTCTCATATCATCTCCATCTCTCCAGATTTGATA 840
 QY 841 CCGCAAAAATTTGTGTTGTGGAGATTTCTGCTGAACTCTTTTCAATTCAGGAAGTA 900
 DB 841 CCGCAAAAATTTGTGTTGTGGAGATTTCTGCTGAACTCTTTTCAATTCAGGAAGTA 900
 QY 901 CCATGATCTGATCTCTGTTGCCCACTGTGCTTAGTAGAATAATTTGGGTGGTACTG 960
 DB 901 CCATGATCTGATCTCTGTTGCCCACTGTGCTTAGTAGAATAATTTGGGTGGTACTG 960

DB 901 CCATGATCTGATCTCTGTTGCCCACTGTGCTTAGTAGAATAATTTGGGTGGTACTG 960
 QY 961 AAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTGATATGT 1020
 DB 961 AAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTGATATGT 1020
 QY 1021 TAAAGTAATGGTGGAGACCAAGAAATTTCAATACAGATCAGTTTGGGGTGCATG 1080
 DB 1021 TAAAGTAATGGTGGAGACCAAGAAATTTCAATACAGATCAGTTTGGGGTGCATG 1080
 QY 1081 TATCATTTTGCAGCTCAATTTGAGTAAGGAAGATTTCTGTATCTTCTGTGCTGGAGGA 1140
 DB 1081 TATCATTTTGCAGCTCAATTTGAGTAAGGAAGATTTCTGTATCTTCTGTGCTGGAGGA 1140
 QY 1141 GGAATGTGTATGATCTCATTTAGATGATCTCCAAAACCTTTTATAAAACCAATTTAGT 1200
 DB 1141 GGAATGTGTATGATCTCATTTAGATGATCTCCAAAACCTTTTATAAAACCAATTTAGT 1200
 QY 1201 TTTAAAAAATAA 1215
 DB 1201 TTTAAAAAATAA 1215
 RESULT 2
 AAI59205
 ID AAI59205 standard; cDNA; 1471 BP.
 XX
 AC AAI59205;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1408.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-USC34263.
 XX
 PR 23-DEC-1999; 99US-03471275.
 PR 21-JAN-2000; 2000US-03488725.
 PR 25-APR-2000; 2000US-02552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSB-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XE, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM40049.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Claim 1; SEQ ID NO 1408; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

58	TTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTTGAGCTATGAGACTGGCTCATTA	117	QY
314	TAAGTAAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTTGAGCTATGAGACTGGCTCATTA	373	Db
118	GAGGCTGTAATGTTGTATACACCAAGTTTCAACGCTCACACCAAGTGAAGACTTCAGAAATTG	177	QY
374	GAGGCTGTAATGTTGTATACACCAAGTTTCAACGCTCACACCAAGTGAAGACTTCAGAAATTG	433	Db
178	AAAACTTTTAAACTTAAATGGTTTATCATATCCAAAAAGACTATCCTCTTAAGTAAGAATT	237	QY
434	AAAACTTTTAAACTTAAATGGTTTATCATATCCAAAAAGACTATCCTCTTAAGTAAGAATT	493	Db
238	TTCCATATTTCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTGTGTCGGAGTTTGATATGC	297	QY
494	TTCCATATTTCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTGTGTCGGAGTTTGATATGC	553	Db
298	GTATGCTTTGCTTAAANAAGCCTTAGGAAATTAGACTTTGAGTCTCAACCAATATAAANAAGC	357	QY
554	GTATGCTTTGCTTAAANAAGCCTTAGGAAATTAGACTTTGAGTCTCAACCAATATAAANAAGC	613	Db
358	TTCCAGCTCAAAATTGGAGACCTCATACACCTTTCAAGAACTTTAACTGAAATGACAAATCACT	417	QY
614	TTCCAGCTCAAAATTGGAGACCTCATACACCTTTCAAGAACTTTAACTGAAATGACAAATCACT	673	Db
418	TGGAGTCAATTTAGTGTAGCCTTGTTGTCACTCTACACTCCAGAACTACTTCGGAGTTTGG	477	QY
674	TGGAGTCAATTTAGTGTAGCCTTGTTGTCACTCTACACTCCAGAACTACTTCGGAGTTTGG	733	Db
478	ACCTCAGCAGAACCAAAANTCAAGCACTCCCTGTGCAGTTTTCGCAGCTCCAGGAACCTTA	537	QY
734	ACCTCAGCAGAACCAAAANTCAAGCACTCCCTGTGCAGTTTTCGCAGCTCCAGGAACCTTA	593	Db
538	AGAATTTTAAACTTGACGATAATCAATTGAATTCAATTTCCCTGCAAGATAGGACAACTAA	597	QY
794	AGAATTTTAAACTTGACGATAATCAATTGAATTCAATTTCCCTGCAAGATAGGACAACTAA	853	Db
598	TAAACCTTCGCTTTTGTGCAGAGCTCGAAATAGCTTCCATTTTTCCTTAGTGAATTTA	657	QY
854	TAAACCTTCGCTTTTGTGCAGAGCTCGAAATAGCTTCCATTTTTCCTTAGTGAATTTA	913	Db
658	GAAATTTATCCCTTGAATATCTTGGAATCTTTTGGAAATCTTTTGGAAACCAACCAAGTCC	717	QY
914	GAAATTTATCCCTTGAATATCTTGGAATCTTTTGGAAATCTTTTGGAAACCAACCAAGTCC	973	Db
718	TTCCAGTAATAAAGCTCGAAGCACCTTAACCTTTTATTTGGAATCTTCTGCAAGCAACCATAT	777	QY
974	TTCCAGTAATAAAGCTCGAAGCACCTTAACCTTTTATTTGGAATCTTCTGCAAGCAACCATAT	1033	Db
778	TACATAATAGGATTCCTCAATGCGGCTCTCAATATCATATCCATTTCCATCTCTGCCAAGATTTGG	837	QY
1034	TACATAATAGGATTCCTCAATGCGGCTCTCAATATCATATCCATTTCCATCTCTGCCAAGATTTGG	1093	Db

CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ Sequence 1471 BP; 433 A; 305 C; 301 G; 432 T; 0 U; 0 Other;
Query Match 94.9%; Score 1153.2; DB 8; Length 1471;
Best Local Similarity 99.7%; Pred. No. 3.7e-290;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTATTCAGGCAATTCAGCAGATTTAAAGGTTTCCCTTTGAGCTATGAGACTGGCTCAT 117
DB 314 TAAGTAAGGCAATTCAGCAGATTTAAAGGTTTCCCTTTGAGCTATGAGACTGGCTCAT 373
QY 118 GAGGCTGTAATGTTGATACACCACTTTCAAGGCTTCAACCACTTCAAGGCTTCAAGGCTT 177
DB 374 GAGGCTGTAATGTTGATACACCACTTTCAAGGCTTCAACCACTTCAAGGCTTCAAGGCTT 433
QY 178 AAAACCTTTAAACCTAAATAGTTATCACTATCAACCAACCAAGACTATCTCTTAAGTAAGATT 237
DB 434 AAAACCTTTAAACCTAAATAGTTATCACTATCAACCAACCAAGACTATCTCTTAAGTAAGATT 493
QY 238 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTGGAGTTGATATGC 297
DB 494 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTGGAGTTGATATGC 553
QY 298 GTATGCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGGTCAACCACTATATAAAGC 357
DB 554 GTATGCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGGTCAACCACTATATAAAGC 613
QY 358 TTCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACT 417
DB 614 TTCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACT 673
QY 418 TGGAGTCAATTTAGTGTAGCTTGTGTATCTTCACTTCCAGAACTCACTTGGAGTTGG 477
DB 674 TGGAGTCAATTTAGTGTAGCTTGTGTATCTTCACTTCCAGAACTCACTTGGAGTTGG 733
QY 478 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGAACTTA 537
DB 734 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGAACTTA 793
QY 538 AGAATTTAAACCTTGAGGATAATGAATGATTAATTTCTTGAAGATAGGACAACTAA 597
DB 794 AGAATTTAAACCTTGAGGATAATGAATGATTAATTTCTTGAAGATAGGACAACTAA 853
QY 598 TAAACCTTCGCTTTTGTGACGAGCTGAAATAGCTTCCATTTTCCCTAGTGAATTTA 657
DB 854 TAAACCTTCGCTTTTGTGACGAGCTGAAATAGCTTCCATTTTCCCTAGTGAATTTA 913
QY 658 GAAATTTATCCCTTGAATCTTGGATCTTTTGAATATCTTTTGAACCAACCAAGTCC 717
DB 914 GAAATTTATCCCTTGAATCTTGGATCTTTTGAATATCTTTTGAACCAACCAAGTCC 973
QY 718 TTCCAGTATTAAGCTGCAAGCACTTAACTTATTTGGAATCTTCTGCAAGCACTAT 777
DB 974 TTCCAGTATTAAGCTGCAAGCACTTAACTTATTTGGAATCTTCTGCAAGCACTAT 1033
QY 778 TACATATAGGATTCATATGGCTCTCATATCATCTTCCATCTCTGCAAGATTTGG 837
DB 1034 TACATATAGGATTCATATGGCTCTCTCATATCATCTTCCATCTCTGCAAGATTTGG 1093
QY 838 ATACCGCAAAATTTGTTTGGAGATTTCTGTCTGAATCTTTCATTTCAAGGAACCTA 897
DB 1094 ATACCGCAAAATTTGTTTGGAGATTTCTGTCTGAATCTTTCATTTCAAGGAACCTA 1153
QY 898 CTACCATGATCTGCATCTCTGTTGCCACACTGTGGTCTTAGTAGATAATTTGGTGGA 957
DB 1154 CTACCATGATCTGCATCTCTGTTGCCACACTGTGGTCTTAGTAGATAATTTGGTGGA 1213
QY 958 CTGAAGCACTTATCTCTTATTTCTTGTCTCTAGGCTGTATGTAATTTCTCTGTATA 1017
DB 1214 CTGAAGCACTTATCTCTTATTTCTTGTCTCTAGGCTGTATGTAATTTCTCTGTATA 1273

QY 1018 TGTAAAGTAATGGTGGAGCAGACAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1077
DB 1274 TGTAAAGTAATGGTGGAGCAGACAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1333
QY 1078 ATGTATGATTTTGGAGCGTCAAAATGGAGTAAGGAAGATTTCTGTATATCTTGGTGGAGA 1137
DB 1334 ATGTATGATTTTGGAGCGTCAAAATGGAGTAAGGAAGATTTCTGTATATCTTGGTGGAGA 1393
QY 1138 GGAGGAATGTATGATGTTACTCAATTTAGATGATCTCCAAACCTTTTATTAACCAATTTT 1197
DB 1394 GGAGGAATGTATGATGTTACTCAATTTAGATGATCTCCAAACCTTTTATTAACCAATTTT 1453
QY 1198 AGTTTAAAAAATAA 1215
DB 1454 AGTTTAAAAAATAA 1471
RESULT 4
ABK90911
ID ABK90911 standard; cdna; 1600 BP.
XX
AC ABK90911;
XX
DT 05-NOV-2002 (first entry)
XX
DE cdna encoding fruit fly LRR47 polypeptide 47-33.88.
XX
KW Fruit fly; LRR47 polypeptide 47-33.88; embryonic development deformity;
XX tumour; diabetes; menstrual disorder; peptide ulcer; arrhythmia; anaemia;
XX epilepsy; gene; ss.
XX
OS Drosophila sp.
XX
FH Key Location/Qualifiers
FT CDS 462..1388
FT /*tag= a
FT /product= "Fruit fly LRR47 polypeptide 47-33.88"
FT
PN CN1341640-A.
XX
XX 27-MAR-2002.
PD
PF 05-SEP-2000; 2000CN-00125025.
XX
XX 05-SEP-2000; 2000CN-00125025.
PR
PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
XX
PI Mao Y, Xie Y;
XX
XX MPI: 2002-520716/56.
DR P-PSDB; ABG31598.
XX
XX A fruit fly LRR47 polypeptide 47-33.88, useful for curing e.g. tumors and
PT diabetes.
XX
PS Claim 6; Page 25-26 (Disclosure); 33pp; Chinese.
XX
CC The present invention relates to a new fruit fly LRR47 polypeptide 47-
CC 33.88. The polypeptide is useful for curing several diseases, such as
CC embryonic development deformity, tumour, diabetes, menstrual disorder,
CC peptide ulcer, arrhythmia, anaemia and epilepsy. The present nucleic acid
CC sequence encodes the fruit fly LRR47 polypeptide 47-33.88 of the
CC invention
XX
SQ Sequence 1600 BP; 455 A; 336 C; 349 G; 460 T; 0 U; 0 Other;
Query Match 94.9%; Score 1153; DB 6; Length 1603;
Best Local Similarity 100.0%; Pred. No. 3.7e-290;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 CAGGCCATTTCCAGCAGTGTAAAGGTTTCCCTTTCAGCTATGAGACTGGCTCATAGAGC 122

Db 423 CAGGCCATTTCCAGCAATTTAAAGGTTTCCTTTAGCTATGAGACTGGCTCATAGGC 482
QY 123 TGTAAATGTTGATACACCAAGTTTCAACGCTTCAACAGTGAAGACTTCAGAAATTTGAAAC 182
Db 483 TGTAAATGTTGATACACCAAGTTTCAACGCTTCAACAGTGAAGACTTCAGAAATTTGAAAC 542
QY 183 TTTAAACCTAAATAGTGTATCATCTCAACCAAAAGAGACTATCTCTAAGTAAAGATTTTCCA 242
Db 543 TTTAAACCTAAATAGTGTATCATCTCAACCAAAAGAGACTATCTCTAAGTAAAGATTTTCCA 602
QY 243 TATTCCTTGGAAATCTTTCAGACTTCTTACTGTGGGCTTTGTCGAGTTGATATGCGTATG 302
Db 603 TATTCCTTGGAAATCTTTCAGACTTCTTACTGTGGGCTTTGTCGAGTTGATATGCGTATG 662
QY 303 CTTTGTCTTAAAGCCTTAGGAAATTAGACTTGTAGTCAACACCNATYATAAAAGCTTCCA 362
Db 663 CTTTGTCTTAAAGCCTTAGGAAATTAGACTTGTAGTCAACACCNATYATAAAAGCTTCCA 722
QY 363 GCTACAAATGGAGACCTCATACACCTTCAAGACTTAACTGAATCAACATCACTTGGAG 422
Db 723 GCTACAAATGGAGACCTCATACACCTTCAAGACTTAACTGAATCAACATCACTTGGAG 782
QY 423 TCATTTAGTGTAGTGTGTGTCTTCTACACTCCAGAGTCACTTCGGAGTTTGGACCTC 482
Db 783 TCATTTAGTGTAGTGTGTGTCTTCTACACTCCAGAGTCACTTCGGAGTTTGGACCTC 842
QY 483 AGCAAGACAAATCAAGGACTCCCTGTGGAGTTTGGAGTTCAGGACTCCAGGAACCTTAAGAAT 542
Db 843 AGCAAGACAAATCAAGGACTCCCTGTGGAGTTTGGAGTTCAGGACTCCAGGAACCTTAAGAAT 902
QY 543 TTAATAACTTGACGATTAATGAATTTCAATTTCCCTTGAAGATAGACAACTTAATAAC 602
Db 903 TTAATAACTTGACGATTAATGAATTTCAATTTCCCTTGAAGATAGACAACTTAATAAC 962
QY 603 CTTCGCTTTTGTGACGAGCTCGAAATAGCTTCCATTTTGGCTAGTGAATTTAGAAAT 662
Db 963 CTTCGCTTTTGTGACGAGCTCGAAATAGCTTCCATTTTGGCTAGTGAATTTAGAAAT 1022
QY 663 TTAATCCCTGATATCTTGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCCCTTCCA 722
Db 1023 TTAATCCCTGATATCTTGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCCCTTCCA 1082
QY 723 GTATAAAGCTCAAGCACCATTAACTTTATTGGAAATCTTTCGACGAACTTATTAACAT 782
Db 1083 GTATAAAGCTCAAGCACCATTAACTTTATTGGAAATCTTTCGACGAACTTATTAACAT 1142
QY 783 AATAGATTCATATGGCTCTCATATCATCTTCATTCATCTCTGCAAGATTTGGATACC 842
Db 1143 AATAGATTCATATGGCTCTCATATCATCTTCATTCATCTCTGCAAGATTTGGATACC 1202
QY 843 GCAAAAATTTGGTTTGGAGATTTCTGCTGAACTTTTCAATCAAGGAATCTATACC 902
Db 1203 GCAAAAATTTGGTTTGGAGATTTCTGCTGAACTTTTCAATCAAGGAATCTATACC 1262
QY 903 ATGAATCTGCATTTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGCTAGAA 962
Db 1263 ATGAATCTGCATTTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGCTAGAA 1322
QY 963 GCACCTATATCTTATTTCTGTCTTAGCTGTATTTGTTAAATTCCTCTGATATGTTA 1022
Db 1323 GCACCTATATCTTATTTCTGTCTTAGCTGTATTTGTTAAATTCCTCTGATATGTTA 1382
QY 1023 AAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATTCAGTTTGGGTGCTAGTA 1082
Db 1383 AAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATTCAGTTTGGGTGCTAGTA 1442
QY 1083 TGATTTTGCAGCGTCAAAATGGAGTAAGGAGATTTCTGTATATCTTGTGAGAGAGG 1142
Db 1443 TGATTTTGCAGCGTCAAAATGGAGTAAGGAGATTTCTGTATATCTTGTGAGAGAGG 1502
QY 1143 AATGTCATAGTTACTCATTTAGATGACTCCAAACTTTTATTAACCAATTTAGTTT 1262

Db 1503 AATGTCATAGTTACTCATTTAGATGACTCCAAACTTTTATTAAACCAATTTAGTTT 1562
QY 1203 TAAAAAATAAAA 1215
Db 1563 TAAAAAATAAAA 1575
RESULT 5
AAS33220
ID AAS33220 standard; cDNA; 1341 BP.
XX AAS33220;
AC AC
XX DT 04-DEC-2001 (first entry)
XX
DE DNA encoding human secreted protein, Seq ID No 179.
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.
XX Homo sapiens.
XX
PN WO20015326-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001347.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 27-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 31-SEP-2000; 2000US-0229287P.
PR 31-SEP-2000; 2000US-0229343P.
PR 31-SEP-2000; 2000US-0229344P.
PR 31-SEP-2000; 2000US-0229345P.

Qy 118 GAGGCTGTAATGTTGATACACCCAGTTTCAACGCTCACACCAGTGAAGACTTCAGAAATTG 177

Db 241 GAGGCTGTATGTTGATACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 300
178 AAAACCTTTAAACTAAATGGTTATACATCCAAAAAGACTATCTCTTAAGTAAGAAAT 237
301 AAAACCTTTAAACTAAATGGTTATACATCCAAAAAGACTATCTCTTAAGTAAGAAAT 360
238 TTCCATATTCCTTGGAGACATCTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATGC 297
361 TTCCATATTCCTTGGAGACATCTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATGC 420
298 GTATGCTTTGCTTTAAAGAGCTTTAGGAANTAGACTTGGATCACAACCATATATAAAGC 357
421 GTATGCTTTGCTTTAAAGAGCTTTAGGAANTAGACTTGGATCACAACCATATATAAAGC 480
358 TTCCAGCTCAATTTGAGAGCTTATACATCTTCAAGAACTTAACTGTAATGCAATCACT 417
481 TTCCAGCTCAATTTGAGAGCTTATACATCTTCAAGAACTTAACTGTAATGCAATCACT 540
418 TGGAGTCATTTAGTGTAGCTTGTGTCACTTCTACACTCCAGAGTCACTTCGGAGTTGG 477
541 TGGAGTCATTTAGTGTAGCTTGTGTCACTTCTACACTCCAGAGTCACTTCGGAGTTGG 600
478 ACTCAGCAGAAACAAAATCAAGGCACTCCCTGTGAGTTTTCAGCTCCAGAACTTA 537
601 ACTCAGCAGAAACAAAATCAAGGCACTCCCTGTGAGTTTTCAGCTCCAGAACTTA 660
538 AGAATTTAAACTTGAAGTATGAAATGATTCATTTCCCTCCAGATAGGCACTAA 597
661 AGAATTTAAACTTGAAGTATGAAATGATTCATTTCCCTCCAGATAGGCACTAA 720
598 TAAACCTTCGCTTTTGTGAGAGCTTGAATTAAGCTTCCATTTTCCCTGAGTGAATTTA 657
721 TAAACCTTCGCTTTTGTGAGAGCTTGAATTAAGCTTCCATTTTCCCTGAGTGAATTTA 780
658 GAAATTTATCCCTTGAATTAAGTGTGATTTTGTGAAATTAAGTGTGAAATTAAGTGTG 717
781 GAAATTTATCCCTTGAATTAAGTGTGATTTTGTGAAATTAAGTGTGAAATTAAGTGTG 840
718 TTCCAGTATTAAGCTTGAAGTATGAAATGATTCATTTCCCTCCAGATAGGCACTAA 777
841 TTCCAGTATTAAGCTTGAAGTATGAAATGATTCATTTCCCTCCAGATAGGCACTAA 899
778 TACATAATAGGATTCATATGCTCTCATATGCTTCCATTCCTCTGCAAGATTTGG 837
900 TACATAATAGGATTCATATGCTCTCATATGCTTCCATTCCTCTGCAAGATTTGG 959
838 ATACGCAAAAATTTGTGTTTGGAGAGATTCCTGTAAGCTTTTCATTCAGGAACTA 897
960 ATACGCAAAAATTTGTGTTTGGAGAGATTCCTGTAAGCTTTTCATTCAGGAACTA 1019
898 STACCAATGATCTGATTCCTGTTGCTCCACACTGCTGTCTTAGTAGATAATTTGGGTGGA 957
1020 STACCAATGATCTGATTCCTGTTGCTCCACACTGCTGTCTTAGTAGATAATTTGGGTGGA 1079
958 CTGAGACCTATTAATCTCTATTTCTGTTCTCTAGGCTGTTAATTTCTCTGATA 1017
1080 CTGAGACCTATTAATCTCTATTTCTGTTCTCTAGGCTGTTAATTTCTCTGATA 1139
1018 TGTATAAGTAAAGGTTGAGACCAAGAAAGAAATTTCAATACAGATCAGTTGGGGTGC 1077
1140 TGTATAAGTAAAGGTTGAGACCAAGAAAGAAATTTCAATACAGATCAGTTGGGGTGC 1199
1078 ATGTATGATTTTGCAGCTCAATTTGGAGTAAGGAGATTTCTGTATATCTCTCTGAGA 1137
1200 ATGTATGATTTTGCAGCTCAATTTGGAGTAAGGAGATTTCTGTATATCTCTCTGAGA 1259
1138 GGAGGAATGTGTATGATTTACTTCAATTTAGATGATCTCCAAATCTTTTATTAACCAATTT 1197
1260 GGAGGAATGTGTATGATTTACTTCAATTTAGATGATCTCCAAATCTTTTATTAACCAATTT 1319
1198 AGTTTTTAAAAA 1214

Db 1320 AGTTTTTAAAAA 1336
RESULT 6
AAI60991
ID AAI60991 standard; cDNA; 1287 BP.
XX
AC AAI60991;
XX
DT 22-OCT-2001 (first entry)
XX
Human polynucleotide SEQ ID NO 4980.
XX
Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
Homo sapiens.
OS
XX
FN WO200153312-A1.
XX
26-JUL-2001.
XX
26-DEC-2000; 2000WO-US034263.
XX
23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
(HYSB-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
WPI; 2001-442253/47.
XX
P-PSDB; NAM41835.
XX
Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
Claim 1; SEQ ID NO 4980; 10078pp; Ect-ish.
XX
The invention relates to human nucleic acids (AAI57798-AAI61369) and the
encoded polypeptides (NAM38642-NAM42213) with nontropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX
SQ Sequence 1287 BP; 363 A; 282 C; 264 G; 378 T; 0 U; 0 Other;
Query Match 78.3%; Score 951.4; DB 4; Length 1287;
Best Local Similarity 99.4%; Pred. No. 1.2e-237;
Matches 955; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 58 TTATTTCAGGCAATTCAGCAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 117
DB 326 TAAGTAAGGCAATTCAGCAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 385
QY 118 GAGGCTGTAATGTTGATACACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 177
DB 386 GAGGCTGTAATGTTGATACACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 445
QY 178 AAAAATTTAAATCAAAATGGTTATCATCATCCAAAAGACTATCCCTTAAGTAAGAAATT 237
DB 446 AAAAATTTAAATCAAAATGGTTATCATCATCCAAAAGACTATCCCTTAAGTAAGAAATT 505
QY 238 TTCCATATTCCTTGGACATCTTCAGACTTCCTACTGTGGCTGTGCTGGAGTTGATATGC 297
DB 506 TTCCATATTCCTTGGACATCTTCAGACTTCCTACTGTGGCTGTGCTGGAGTTGATATGC 565
QY 298 GTATGCTTGTCTTAAAGGCTTTAGGAAATTAGACTTGGTCAACCATATATAAAGC 357
DB 566 GTATGCTTGTCTTAAAGGCTTTAGGAAATTAGACTTGGTCAACCATATATAAAGC 625
QY 358 TTCCAGCTACAATGGAGACTCATACCTTCAAGAACTTAACTGAATGACAACTACT 417
DB 626 TTCCAGCTACAATGGAGACTCATACCTTCAAGAACTTAACTGAATGACAACTACT 685
QY 418 TGGAGTCAATTAAGTGTAGCTGTGCTTCACTTCTACACTCCAGAACTCACTCGGAGTTGG 477
DB 686 TGGAGTCAATTAAGTGTAGCTGTGCTTCACTTCTACACTCCAGAACTCACTTTGGAGTTGG 745
QY 478 ACCTCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGAACTTA 537
DB 746 ACCTCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGAACTTA 805
QY 538 AGAATTTAAACTGTGACGTAATGAATGAATCAATTCCTTGCAGATAGGCAACTAA 597
DB 806 AGAATTTAAACTGTGACGTAATGAATGAATCAATTCCTTGCAGATAGGCAACTAA 865
QY 598 TAAACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCATTTTGCCTAGTGAATTA 657
DB 866 TAAACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCATTTTGCCTAGTGAATTA 925
QY 658 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCC 717
DB 926 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCC 985
QY 718 TTCCAGTAATTAAGCTGCAAGCAACCAATTAATTTTGAATCTTTCGACGCAACCATAT 777
DB 986 TTCCAGTAATTAAGCTGCAAGCAACCAATTAATTTTGAATCTTTCGACGCAACCATAT 1045
QY 778 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATTCATTCATTCATTCATTCAT 837
DB 1045 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATTCATTCATTCATTCATTCAT 1105
QY 838 ATACCCGCAAAATTTGTTTGGAGAGATTCGTCTGAATCTTTTCATTCAGGAACTA 897
DB 1106 ATACCCGCAAAATTTGTTTGGAGAGATTCGTCTGAATCTTTTCATTCAGGAACTA 1165
QY 898 CTACCATGATCTGCAATCTGTTGCCACACTGCTGTCTAGTAGATAAATTCGGTGGTA 957
DB 1166 CTACCATGATCTGCAATCTGTTGCCACACTGCTGTCTAGTAGATAAATTCGGTGGTA 1225
QY 958 CTGAAGCACTTATCTCTTATTTCTGTTCTCTAGGCTGTATGTAATTCCTCTGATA 1017
DB 1226 CTGAAGCACTTATCTCTTATTTCTGTTCTCTAGGCTGTATGTAATTCCTCTGATA 1285
QY 1018 T 1018
DB 1286 T 1286

RESULT 7
AAD27818
ID AAD27818 standard; DNA; 1251 BP.

XX AAD27818;
AC 18-APR-2002 (first entry)
DT Human leucine rich repeat protein gene, sbg442358PROA #2.
XX
DE
XX
XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
KW depression; cardiovascular disease; myocardial infarction; renal failure;
KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
KW hyperplasia; renal disease; hypoglycaemia; gastrointestinal disease;
KW neoplastic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
KW haemostatic; vulvovaginitis; anticonvulsant; antirheumatic; neuroprotective;
KW nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;
KW allergy; leucine rich repeat protein; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1.1251
FT /tag= a
FT /product= "Human leucine rich repeat protein"
XX WO200198342-A1.
XX
XX 27-DEC-2001.
XX
XX 22-JUN-2001; 2001WO-US019929.
XX
XX 22-JUN-2000; 2000US-0213156P.
XX
XX 22-JUN-2000; 2000US-0213161P.
XX
XX (SWIK) SMITHKLINE BEECHAM CORP.
XX (SWIK) SMITHKLINE BEECHAM PLC.
XX (GLAX) GLAXO GROUP LTD.
XX
XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA,
PI Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;
XX
XX WPI: 2002-139783/18.
XX P-PSDB; AAE17323.
XX
XX Novel secreted and membrane-associated polypeptides and polynucleotides
XX useful for preventing, ameliorating or correcting dysfunction or disease
XX including diabetes, cancer, hypertension and growth abnormalities.
XX
XX Claim 2; Page 99; 138pp; English.
XX
XX The invention relates to secreted and membrane-associated polypeptides
XX and polynucleotides. The sequences of the invention are useful in
XX diagnostic assays for detecting diseases associated with inappropriate
XX activity or levels of these polynucleotides, and in identifying their
XX agonists and antagonists that are potentially useful in therapy. The
XX sequences of the invention are useful as vaccines for inducing
XX immunological response. The sequences of the invention are useful for
XX treating cancers, infections, autoimmune disorders, haematopoietic
XX disorders, wound healing disorders, cholesterol ester storage disease,
XX inflammation, congenital muscular dystrophy, junctional epidermolysis
XX bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,
XX viral and bacterial infections, Alzheimer's disease, asthma, arthritis,
XX allergies, schizophrenia, sbg44245PROA-associated disorders,
XX septicemia, psoriasis, inflammatory bowel disease, transplant rejection,
XX graft versus host disease, ischaemia, stroke, acute respiratory disease
XX syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis,
XX brain disorders including parasupranuclear palsy, myotonic dystrophy,
XX depression, anxiety disorders and sleep disorders, cardiovascular
XX diseases including congestive heart failure and myocardial infarction,
XX respiratory diseases including chronic obstructive pulmonary disease,

CC acute bronchitis and adult respiratory distress syndrome, liver disorders
CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral
CC and non-viral hepatitis, type II diabetes mellitus, renal disease
CC syndrome, acute and chronic renal failure, glomerulonephritis, Fanconi's
CC disease, cystinuria, skeletal muscle disorders including hypoglycaemia
CC and tendinitis, gastrointestinal diseases including intestinal
CC obstruction and tropical sprue, spleen disorders including hypersplenism,
CC Hodgkin's disease and malignant lymphoma, testicular cancer, male
CC reproductive diseases including low testosterone and male infertility.
CC The present sequence is human leucine rich repeat protein gene
XX
SQ Sequence 1251 BP; 358 A; 275 C; 248 G; 370 T; 0 U; 0 Other;

Query Match 78.2%; Score 950.2; DB 6; Length 1251;
Best Local Similarity 99.1%; Pred. No. 2.4e-237;
Matches 968; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 58 TTATTCAGGCCAATTCACGAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 117
DB 275 TAAGTAAGGCCAATTCACGAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 334
QY 118 GAGGCTGTAATGTGATACACAGTTTCAACGCTCACACAGTGRAGACTTCAGAAATTG 177
DB 335 GAGGCTGTAATGTGATACACAGTTTCAACGCTCACACAGTGRAGACTTCAGAAATTG 394
QY 178 AAAAAGCTTTAAAGCTAAATGGTTATCACATCCAAAAGAGACTATCCTCTAAGTAAGAATT 237
DB 395 AAAAGCTTTAAAGCTAAATGGTTATCACATCCAAAAGAGACTATCCTCTAAGTAAGAATT 454
QY 238 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTTCGAGTTGATATGC 297
DB 455 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTTCGAGTTGATATGC 514
QY 298 GTATGCTTTGCTTAAAGAGCTTAGGAATTTAGACTGTGACTCACACCAATATAAAGAGC 357
DB 515 GTATGCTTTGCTTAAAGAGCTTAGGAATTTAGACTGTGACTCACACCAATATAAAGAGC 574
QY 358 TTCCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAAATCACT 417
DB 575 TTCCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAAATCACT 634
QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCATTTACACTCCAGAAAGTCACTTCGAGTTTGG 477
DB 635 TGGAGTCATTTAGTGTAGCTTGTGTCATTTACACTCCAGAAAGTCACTTCGAGTTTGG 694
QY 478 ACCTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCAGTTTTCGCACTCCAGAACTTA 537
DB 695 ACCTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCAGTTTTCGCACTCCAGAACTTA 754
QY 538 AGAATTTAAAGCTTGAAGATAATGAATGATCAATTTCTTCGAAGATAGGACAACTAA 597
DB 755 AGAATTTAAAGCTTGAAGATAATGAATGATCAATTTCTTCGAAGATAGGACAACTAA 814
QY 598 TAAACCTTCGCTTTTGTGAGAGCTCGAATAGCTTCCATTTTTCCTAGTGAATTTA 657
DB 815 TAAACCTTCGCTTTTGTGAGAGCTCGAATAGCTTCCATTTTTCCTAGTGAATTTA 874
QY 658 GAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATATCTTTTGAAACCAACAAAGTCC 717
DB 875 GAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATATCTTTTGAAACCAACAAAGTCC 934
QY 718 TTCCAGTAATAAGCTGCAAGCAACCAATTAATCTTATTTGAATCTTCGACGCAACCATAT 777
DB 935 TTCCAGTAATAAGCTGCAAGCAACCAATTAATCTTATTTGAATCTTCGACGCAACCATAT 994
QY 778 TACAT-----AATAGGATTCATATGCTTCATATCAATTCATCTCTGCGAAG 831
DB 995 TACATTAATAGGAATAGGATTCATATGCTTCATATCAATTCATCTCTGCGAAG 1054
QY 832 ATTTGGATACCGCAAAATTTGTGTTGTGGAGATTTCTGTCTGAATCTTTTCATTTCAAG 891
DB 1055 ATTTGGATACCGCAAAATTTGTGTTGTGGAGATTTCTGTCTGAATCTTTTCATTTCAAG 1114

QY 892 GAACCTACTACCAATCTGATCTCTGTTGCCACACACTGTGGTCTTAGTAGATAATTTGG 951
DB 1115 GAACCTACTACCAATCTGATCTCTGTTGCCACACACTGTGGTCTTAGTAGATAATTTGG 1174
QY 952 GTCGTACTAGAACCACTTATCTTATTTCTGTTCTTCTAGGCTGTATGTTAATTCCT 1011
DB 1175 GTCGTACTAGAACCACTTATCTTATTTCTGTTCTTCTAGGCTGTATGTTAATTCCT 1234
QY 1012 CTGATATCTTAAAGTAA 1028
DB 1235 CTGATATCTTAAAGTAA 1251

RESULT 8
AAD27817

ID AAD27817 standard; DNA; 780 BP.

XX

XX AAD27817;

XX 18-APR-2002 (first entry)

XX Human leucine rich repeat protein gene, sbg442358PROa #1.

XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
XX autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
XX Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
XX multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
XX ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
XX depression; cardiovascular disease; myocardial infarction; renal failure;
XX respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
XX type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
XX hypersplenism; renal disease; hypoglycaemia; gastrointestinal disease;
XX neutropenic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
XX haemostatic; vulvular; anticonvulsant; antineoplastic; neuroprotective;
XX nephrotropic; hypotensive; vasotropic; cycostatic; cerebroprotective;
XX allergy; leucine rich repeat protein; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..780

XX FT /tag= a

XX FT /product= "Human leucine rich repeat protein"

XX PN WO200198342-A1.

XX PD 27-DEC-2001.

XX PF 22-JUN-2001; 2001WO-US019929.

XX PR 22-JUN-2000; 2000US-0213156P.

XX PR 22-JUN-2000; 2000US-0213161P.

XX XX (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PA (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;

XX Yurdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;

XX DR P-PSDB; AAE17322.

XX Novel secreted and membrane-associated polypeptides and polynucleotides

XX useful for preventing, ameliorating or correcting dysfunction or disease

XX including diabetes, cancer, hypertension and growth abnormalities.

XX Claim 2; Page 98; 138pp; English.

XX The invention relates to secreted and membrane-associated polypeptides

XX and polynucleotides. The sequences of the invention are useful in

XX diagnostic assays for detecting diseases associated with inappropriate

XX activity or levels of these polynucleotides, and in identifying their


```
CC the invention
XX
SQ Sequence 682 BP; 200 A; 140 C; 131 G; 200 T; 0 U; 11 Other;

Query Match      42.9%; Score 521.6; DB 4; Length 682;
Best Local Similarity 97.5%; Pred. No. 9.9e-126;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY 277 GCGTGTGCGAGTTGATATGCGTATGCTTTGGCTTTAAAGAGCCTTAGGAATTAGACTTGA 336
   |||
Db 1 GCGTGTGCGAGTTGATATGCGTATGCTTTGGCTTTAAAGAGCCTTAGGAATTAGACTTGA 60

QY 337 GTCACCAACATATAAAAAAGCTTCCAGCTACATTTGGAGACCTTCATACACCTTCAAGAAC 396
   |||
Db 61 GTCACCAACATATAAAAAAGCTTCCAGCTACATTTGGAGACCTTCATACACCTTCAAGAAC 120

QY 397 TTAACCTGAATGACATCTCTGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACACTCC 456
   |||
Db 121 TTAACCTGAATGACATCTCTGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACACTCC 180

QY 457 AGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCGAGT 516
   |||
Db 181 AGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCGAGT 240

QY 517 TTGCGAGCTCCAGGAACCTTAAGAATTTAAAGCTTGACGATATGAATGATTCATTTTC 576
   |||
Db 241 TTGCGAGCTCCAGGAACCTTAAGAATTTAAAGCTTGACGATATGAATGATTCATTTTC 300

QY 577 CTTGCAAGTAGACACCTTAATTAACCTTCGCTTTTGTGACGAGCTCGAAATAGCTTC 636
   |||
Db 301 CTTGCAAGTAGACACCTTAATTAACCTTCGCTTTTGTGACGAGCTCGAAATAGCTTC 360

QY 637 CATTTTGTCTAGTCAATTTAGAAATTTATCCCTTGAATACCTTTTGTGAAATA 696
   |||
Db 361 CATTTTGTCTAGTCAATTTAGAAATTTATCCCTTGAATACCTTTTGTGAAATA 420

QY 697 CTTTGTGAACACCAAAAGCTTCCAGTAAATAAGCTGCAAGCACCATTATTTATTCG 756
   |||
Db 421 CTTTGTGAACACCAAAAGCTTCCAGTAAATAAGCTGCAAGCACCATTATTTATTCG 480

QY 757 AATCTTCTGACGACCATATTACATATATAGATTCATATGCTCT---CATATCAT 812
   |||
Db 481 AATCTTCTGACGACCATATTACATATATAGATTCATATGCTCTCTCATATTCATTT 540

QY 813 CCATTCATCTCTGCCAAGATTTGG 837
   |||
Db 541 CCATTCATCTCTGCCAAGATTTGG 565
```

```
RESULT 20
ABL68766/c
ID ABL68766 standard; DNA; 467 BP.
XX
AC ABL68766;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:7103.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; anticarcinoma; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
CS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US0-0838.
XX
XX 05-JUN-2000; 2000US-0209473P.
XX
```

05-JUN-2000; 2000US-0209531P.
18-SEP-2000; 2000US-0233333P.
18-SEP-2000; 2000US-0233617P.
20-SEP-2000; 2000US-0234009P.
20-SEP-2000; 2000US-0234034P.
20-SEP-2000; 2000US-0234052P.
22-SEP-2000; 2000US-0234509P.
22-SEP-2000; 2000US-0234567P.
25-SEP-2000; 2000US-0234523P.
25-SEP-2000; 2000US-0234924P.
25-SEP-2000; 2000US-0235077P.
25-SEP-2000; 2000US-0235082P.
25-SEP-2000; 2000US-0235134P.
25-SEP-2000; 2000US-0235280P.
26-SEP-2000; 2000US-0235637P.
26-SEP-2000; 2000US-0235638P.
27-SEP-2000; 2000US-0235711P.
27-SEP-2000; 2000US-0235720P.
27-SEP-2000; 2000US-0235840P.
27-SEP-2000; 2000US-0235863P.
28-SEP-2000; 2000US-0236032P.
28-SEP-2000; 2000US-0236033P.
28-SEP-2000; 2000US-0236034P.
28-SEP-2000; 2000US-0236109P.
28-SEP-2000; 2000US-0236111P.
29-SEP-2000; 2000US-0236842P.
29-SEP-2000; 2000US-0236891P.
02-OCT-2000; 2000US-0237172P.
02-OCT-2000; 2000US-0237173P.
02-OCT-2000; 2000US-0237278P.
02-OCT-2000; 2000US-0237294P.
02-OCT-2000; 2000US-0237295P.
02-OCT-2000; 2000US-0237316P.
03-OCT-2000; 2000US-0237425P.
03-OCT-2000; 2000US-0237598P.
03-OCT-2000; 2000US-0237604P.
03-OCT-2000; 2000US-0237606P.
03-OCT-2000; 2000US-0237608P.
01-NOV-2000; 2000US-0244867P.
01-NOV-2000; 2000US-0244867P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI; 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.
Claim 1; SEQ ID NO 7103; 44pp; English.
The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour

QY 1206 AAAAAAAAAA 1215
 DB 1180 ACAAACACA 1189

RESULT 12
 ID AAH09329/c
 ID AAH09329 standard; cDNA; 584 BP.
 AC AAH09329;
 XX
 XX
 XX 26-JUN-2001 (first entry)
 XX
 XX Human cDNA clone (3'-primer) SEQ ID NO:6164.
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW Homo sapiens.
 OS
 XX
 XX EP1074617-A2.
 XX
 XX C7-FEB-2001.
 PD
 XX
 XX 28-JUL-2000; 2000EP-00116126.
 XX
 XX 29-JUL-1999; 99JP-00248036.
 PR
 XX 27-AUG-1999; 99JP-00300253.
 PR
 XX 11-JAN-2000; 2000JP-00118776.
 PR
 XX 02-MAY-2000; 2000JP-00183767.
 PR
 XX 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX
 XX Claim 3; SEQ ID NO 6164; 2537pp + Sequence Listing; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX
 XX Sequence 584 BP; 177 A; 132 C; 121 G; 146 T; 0 U; 8 Other;
 SQ
 Query Match 34.5%; Score 418.9; DB 4; Length 584;
 Best Local Similarity 99.5%; Pred. No. 5.8e-99;
 Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 782 TATAGATTCCATATGGCTCTCATATCATTTCCATTCCTGCGCAAGATTGGATAC 841
 DB 422 TATAGGATTCCATATGGCTCTCATATCATTTCCATTCCTGCGCAAGATTGGATAC 363
 QY 842 CGCAAAAATTGTGTTGTGGAAGATTCTCTCTGAACTCTTTTCAATTCAGGAACTACTAC 901
 DB 362 CGCAAAAATTGTGTTGTGGAAGATTCTCTCTGAACTCTTTTCAATTCAGGAACTACTAC 303
 QY 902 CATGAATCTGCATTCCTGTTGCCACACCTGTGGTCTTAGTAGATAAATTTGGTGGTACTCGA 961
 DB 302 CATGAATCTGCATTCCTGTTGCCACACCTGTGGTCTTAGTAGATAAATTTGGTGGTACTCGA 243
 QY 962 AGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAATTCCTCTCATATGTT 1021
 DB 242 AGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAATTCCTCTCATATGTT 183
 QY 1022 AAGTAATGGTGGACACAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATGT 1081
 DB 182 AAGTAATGGTGGACACAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATGT 123
 QY 1082 ATGATTTTGCAGGCTCAAAATTTGGAGTAAGGAAAGATTTCTGTATACCTGTGGAGAGGAG 1141
 DB 122 ATGATTTTGCAGGCTCAAAATTTGGAGTAAGGAAAGATTTCTGTATACCTGTGGAGAGGAG 63
 QY 1142 GAATGTATAGTACTCTCAATTTAGATGACTCCAAAACCTTTTATTAACCAATTTAGTT 1201
 DB 62 GAATGTATAGTACTCTCAATTTAGATGACTCCAAAACCTTTTATTAACCAATTTAGTT 3
 QY 1202 TT 1203
 DB 2 TT 1

RESULT 13
 AAH15353
 ID AAH15353 standard; cDNA; 1590 BP.
 XX
 AC AAH15353;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:13527.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD C7-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 DR
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 PT
 XX
 PS Claim 8; SEQ ID NO 13527; 2537pp + Sequence Listing; English.

AA98737

ID AAX98737 standard; cDNA; 750 BP.

XX AAX98737;

XX 24-SEP-1999 (first entry)

XX Human validated cancer cell derived cDNA #59.

XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
 KW integral membrane protein; aspartyl protease; GATA family; wnt family;
 KW transcription factor; G-protein alpha subunit; protein phosphatase;
 KW phorbol ester binding protein; diacylglycerol binding protein; trypsin;
 KW protein kinase; tyrosine phosphatase; developmental signalling protein;
 KW WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;
 KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
 KW Wilms' tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
 KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
 KW prostate; ss.

XX Homo sapiens.

XX OS

XX WO9933982-A2.

XX PN

XX 08-JUL-1999.

XX PD

XX 22-DEC-1998; 98WO-US027610.

XX PF

XX 23-DEC-1997; 97US-0068755P.

XX PR 03-APR-1998; 98US-0080664P.

XX PR 21-OCT-1998; 98US-0105234P.

XX PR 27-OCT-1998; 98US-0105877P.

XX PR 21-DEC-1998; 98US-00217471.

XX XX

XX (CHIR) CHIRON CORP.

XX PA (HYSE-) HYSEQ INC.

XX XX

XX Williams LT, Escobedo J, Irmis MA, Garcia PD, Sudduth-Klinger J;

XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

XX Larson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

XX PI Leskowitz D, Kita D, Garcia V, Jones LM, Stache-Crain B;

XX XX

XX WPI; 1999-430243/36.

XX DR

XX New isolated human polynucleotides.

XX PT

XX Claim 1; Page 437; 591pp; English.

XX PS

XX This invention describes novel isolated human polynucleotides obtained by

XX screening for differential expression in colon cancer, breast cancer and

XX lung cancer cell lines. The polynucleotides of the invention are

XX represented in AAX98275-X99118 and encode polypeptides of protein

XX families selected from 4 transmembrane segments integral membrane

XX proteins, 7 transmembrane receptors, ATPases associated with various

XX cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of

XX transcription factors, G-protein alpha subunit, phorbol esters or

XX diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,

XX protein tyrosine phosphatase, trypsin, wnt family of developmental

XX signalling proteins and WW/rsp5/WWP domain containing proteins. The

XX encoded polypeptides also have a functional domain selected from Ank

XX repeat, basic region plus leucine zipper transcription factors,

XX bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger

XX (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease

XX domain. The polynucleotides encode polypeptides with similarity to known

XX protein families and are predicted to have similar properties. The novel

XX polynucleotides can be used to develop products for use as therapeutic

XX agents and in forensics, genetic analysis, mapping and diagnostic

XX applications. In particular, the product can be used for the detection

XX and management of cancers. They can be used for treating e.g. cervical

XX cancers, melanomas, colorectal adenocarcinomas, Wilms' tumour, sarcomas,

XX retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic

XX myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and

XX myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric

CC hereditary ectodermal dysplasia, congenital alveolar dysplasia,
 CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and
 CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,
 CC prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of
 CC the skin

XX SQ Sequence 750 BP; 184 A; 158 C; 179 G; 187 T; 0 U; 42 Other;

Query Match 23.2%; Score 281.4; DB 2; Length 750;

Best Local Similarity 87.0%; Pred. No. 4.1e-63;

Matches 320; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

QY 786 AGGATTCATATGGCTCTCATATCCATTCATTCATCTCTGCCAAGATTTGGATACCGCA 845

Db 383 AGGATTCATATGGCTCTCATATCCATTCATTCATCTCTGCCAAGATTTGGATACCGCA 442

QY 846 AAAAATTTGTTGTTGGAAGATTTCTGTCTGAATCTTTTCAATTCAGGAATCTACTPACCATG 905

Db 443 AAAAATTTGTTGTTGGAAGATTTCTGTCTGAATCTTTTCAATTCAGGAATCTACTPACCATG 502

QY 906 AATCTGCATCTCTTGGCCACACTGTGCTCTTAGTAGATAAATTTGGTGGTACTGAGCA 965

Db 503 AATCTGCATCTCTTGGCCACACTGTGCTCTTAGTAGATAAATTTGGTGGTACTGAGCA 560

QY 966 CCTATTATCTCTTATTTCTGTTCTAGGCTGTTATGTTAAATTCCTCTGATATGTT-AAA 1024

Db 561 CCTATTATCTCTTATTTCTGTTCTAGGCTGTTATGTTAAATTCCTCTGATATGTT-AAA 620

QY 1025 GTAATGGGTGAGACCCAGAAAAGAAATTTTCAATAACAGATCAGTTTGGGGTGCATGTATG 1084

Db 621 GTAATGGGTGAGACCCAGAAAAGAAATTTTCAATAACAGATCAGTTTGGGGTGCATGTATG 680

QY 1085 ATTTTCAGCGTCAAATTTGGAGTAAGGGAAGATTTCTGTATACCTTGTGGAGAGAGGAA 1144

Db 681 ATTTTCAGCGTCAAATTTGGAGTAAGGGAAGATTTCTGTATACCTTGTGGAGAGAGGAA 740

QY 1145 TGTGTATA 1152

Db 741 NATGTGTA 748

Search completed: June 15, 2004, 23:20:52

Job time : 403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 21:45:13 ; Search time 3273 Seconds
(without alignments)
16089.761 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagcatgtatgta.....ttagttaaataaaaaa 1215

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.scs.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rtd.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1215	100.0	1215	6	BD276293	BD276293 EXTRACELL
2	1215	100.0	1215	6	AX047348	AX047348 Sequence
3	1153.2	94.9	1471	6	AR339608	AR339608 Sequence
4	1153.2	94.9	1613	9	EC064133	EC064133 Homo sapi
5	1138	93.7	1445	9	AY052405	AY052405 Homo sapi
6	986.8	81.2	1950	9	BSM808175	BSM808175 Homo sapi
7	788.6	64.9	157010	9	CNS01DX7	AL139099 Human chr
8	565.6	46.8	193800	2	AC126994	AC126994 Rattus no
9	566	46.6	232458	2	AC099934	AC099934 Mus muscu
10	521.6	42.9	682	6	AX255882	AX255882 Sequence
11	443.8	36.5	467	6	AX336594	AX336594 Sequence
12	430	35.4	990	9	BC030142	BC030142 Homo sapi
13	418.8	34.5	584	6	AX871259	AX871259 Sequence
14	418.8	34.5	584	6	BD151321	BD151321 Primer fo
15	418.8	34.5	1590	6	AX878622	AX878622 Sequence
16	418.8	34.5	1590	6	BD157345	BD157345 Primer fo
17	418.8	34.5	1590	9	AK027315	AK027315 Homo sapi
18	416.2	34.3	75921	2	AC107185	Continuation (4 of
19	387	31.9	1681	5	BC064158	BC064158 Silurana
20	281.4	23.2	750	6	BD129164	BD129164 Human gen
21	281	23.1	732	6	BD129168	BD129168 Human gen
22	212.8	17.5	751	10	BC022648	BC022648 Mus muscu
23	188.8	15.5	770	6	BD215981	BD215981 Novel hum
24	115.4	9.5	132396	9	AL607023	AL607023 Human DNA
25	91.6	7.5	430	6	AX070036	AX070036 Sequence
26	85.2	7.0	382	6	AX070144	AX070144 Sequence
27	84.4	6.9	236640	2	AC114464	AC114464 Rattus no
28	84.4	6.9	270622	2	AC115235	AC115235 Rattus no
29	66	5.4	227056	2	AC121316	AC121316 Mus muscu
30	60.2	5.0	2950	6	AX552112	AX552112 Sequence
31	59.6	4.9	1658	9	AF359380	AF359380 Homo sapi
32	59.6	4.9	2156	9	BC003193	BC003193 Homo sapi
33	59.6	4.9	2886	9	AF332199	AF332199 Homo sapi
34	56.2	4.6	2169	6	BD251848	BD251848 RING fing
35	56.2	4.6	2169	6	AR224072	AR224072 Sequence
36	55	4.5	4941	9	BSM807337	BSM807337 Homo sapi
37	55	4.5	4984	6	AX556708	AX556708 Sequence
38	54.4	4.5	2694	6	BD232164	BD232164 A novel m
39	53.8	4.4	11990	1	AE010918	AE010918 Methanosa
40	53.8	4.4	164026	2	AC092762	AC092762 Pan trogl
41	53.8	4.4	173456	2	AC092763	AC092763 Pan trogl
42	53.4	4.4	1141	6	AX083744	AX083744 Sequence
43	53.4	4.4	128833	2	AC135793	AC135793 Oryza sat
44	52.4	4.3	2620	6	AX877169	AX877169 Sequence
45	52.4	4.3	2620	6	BD156521	BD156521 Primer fo

ALIGNMENTS

RESULT 1
BD276293
LOCUS BD276293
DEFINITION EXTRACELLULAR MATRIX AND ADHESION-ASSOCIATED PROTEINS.
ACCESSION BD276293
VERSION BD276293.1 GI:33086061
KEYWORDS JP 2002543785-A/9.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1215)
AUTHORS Azimzai, Y., Bandman, O., Tang, T.Y., La, P., Henry, Yue, Baughn, M.R.,
Lu, D.A.M. and Hillman, J.L.
TITLE EXTRACELLULAR MATRIX AND ADHESION-ASSOCIATED PROTEINS

Pred. No. is the number of results predicted by chance to have a

JOURNAL

Patent: JP 2002543785-A 9 24-DEC-2002;
 INCYTE PHARMACEUTICALS INC, Olga BANDMAN, Jennifer L HILLMAN, Tom Y
 TANG, Preeti Lal, Henry YUE, Maria R BAUGHN, Dyung Aina M LU, Yalda
 AZIMZAI

COMMENT

OS Homo sapiens
 PN JP 2002543785-A/9
 PD 24-DEC-2002
 PF 10-MAY-2000 JP 2000616346
 PR 11-MAY-1999 US 60/13643.23-AUG-1999 US 60/150409 PI
 Yalda azimzai, Olga bandman, tom y tang, preeti lal, henry pi
 yue, maria r baughn,
 PI dyung aina m lu, jennifer l hillman
 CC This description about <220> can't be interpreted CC <220>
 CC <221> misc feature
 CC <223> Incyte ID No: 2040679CB1

FEATURES

Key Location/Qualifiers
 1. .1215
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1215; DB 6; Length 1215;
 Best Local Similarity 100.0%; Pred. No. 1.8e-261;
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGAACTAGCATGTATGTTATCTCCAGTGGAAATTTAAATCTCAACATTTTATTA 60
 DB 1 GAAGAACTAGCATGTATGTTATCTCCAGTGGAAATTTAAATCTCAACATTTTATTA 60
 QY 61 TTGAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGCTCATAGAG 120
 DB 61 TTGAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGCTCATAGAG 120
 QY 121 GCTGTAAATGTTGATACACCACTTCAACGCTCTACACAGTGAAGACTTCAAGATTTGAAA 180
 DB 121 GCTGTAAATGTTGATACACCACTTCAACGCTCTACACAGTGAAGACTTCAAGATTTGAAA 180
 QY 181 ACTTTAAACTAAATGGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGATTTTC 240
 DB 181 ACTTTAAACTAAATGGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGATTTTC 240
 QY 241 CATATTCCTTGGAACTCTTCACTCTTCTTCTGAGCTTCTGAGTGTGATATGCGTA 300
 DB 241 CATATTCCTTGGAACTCTTCACTCTTCTTCTGAGCTTCTGAGTGTGATATGCGTA 300
 QY 301 TGCCTTGTCTTAAAGGCTTAGGAATTTAGACTTGTGATCAACACCATATAAAAAGCTTC 360
 DB 301 TGCCTTGTCTTAAAGGCTTAGGAATTTAGACTTGTGATCAACACCATATAAAAAGCTTC 360
 QY 361 CAGCTACAACTGGAGACCTCATACACCTTCAAGAACTTAACTGGAATGCAATCACTTGG 420
 DB 361 CAGCTACAACTGGAGACCTCATACACCTTCAAGAACTTAACTGGAATGCAATCACTTGG 420
 QY 421 AGTCATTTAGTGTAGCTTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
 DB 421 AGTCATTTAGTGTAGCTTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
 QY 481 TCAGCAAGAACAAATCAAGGCACTCCCTGTGAGTTTCCAGTTCAGGAACTTAAGA 540
 DB 481 TCAGCAAGAACAAATCAAGGCACTCCCTGTGAGTTTCCAGTTCAGGAACTTAAGA 540
 QY 541 ATTTAAACTTGACGATATGATTTGATTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 600
 DB 541 ATTTAAACTTGACGATATGATTTGATTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 600
 QY 601 ACCTTCGCTTTTGTGACGAGCTCGAAATTAAGTTCCTATTTTGTGCTAGTGAATTTAGAA 660
 DB 601 ACCTTCGCTTTTGTGACGAGCTCGAAATTAAGTTCCTATTTTGTGCTAGTGAATTTAGAA 660
 QY 661 ATTTATCCCTTGAATCTTGGATCTTTTGTGAAATTAATTTTCAACCAACCAAGTCTTTC 720
 DB 661 ATTTATCCCTTGAATCTTGGATCTTTTGTGAAATTAATTTTCAACCAACCAAGTCTTTC 720

Db 661 ATTTATCCCTTGAATCTTGGATCTTTTGTGAAATTAATTTTCAACCAACCAAGTCTTTC 720
 QY 721 CAGTAATAAAGCTGCAAGCACCACCAATTAATTTTGAATCTTCTGCAAGAACCAATATTAC 780
 Db 721 CAGTAATAAAGCTGCAAGCACCACCAATTAATTTTGAATCTTCTGCAAGAACCAATATTAC 780
 QY 781 ATATAGAGTTCCATATGAGCTCTCATATCCATTCATTCATTCATTCATTCATTCATTCATTC 840
 Db 781 ATATAGAGTTCCATATGAGCTCTCATATCCATTCATTCATTCATTCATTCATTCATTCATTC 840
 QY 841 CCGCAAAATTTTGTGGAAGACTTCTGCAAGCTTCTGCAAGCTTCTGCAAGCTTCTGCAAGCTTCT 900
 Db 841 CCGCAAAATTTTGTGGAAGACTTCTGCAAGCTTCTGCAAGCTTCTGCAAGCTTCTGCAAGCTTCT 900
 QY 901 CCATGAATCTGCAATCTGTTGCCACACTGTTGCTTCTAGTAGAATTTGCTGCTGCTGCTGCTG 960
 Db 901 CCATGAATCTGCAATCTGTTGCCACACTGTTGCTTCTAGTAGAATTTGCTGCTGCTGCTGCTG 960
 QY 961 AAGCACCCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
 Db 961 AAGCACCCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
 QY 1021 TAAAGTAATGGTGGAGACCAAGAAATTTCAATTAACAGATCAGTTTGGGCTGCAATG 1080
 Db 1021 TAAAGTAATGGTGGAGACCAAGAAATTTCAATTAACAGATCAGTTTGGGCTGCAATG 1080
 QY 1081 TATGATTTTGGAGCTCAAAATTTGAGTAAGGAAATTTCTGTATATCTTCTGCAAGAGGA 1140
 Db 1081 TATGATTTTGGAGCTCAAAATTTGAGTAAGGAAATTTCTGTATATCTTCTGCAAGAGGA 1140
 QY 1141 GGAATGTTAGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
 Db 1141 GGAATGTTAGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
 QY 1201 TTTAAAAAATTTTAAAAA 1215
 Db 1201 TTTAAAAAATTTTAAAAA 1215

RESULT 2
 AX047348 1215 bp DNA linear PAT 15-DEC-2000
 LOCUS Sequence 34 from Patent WO0068380.
 DEFINITION AX047348
 ACCESSION AX047348
 VERSION AX047348.1 GI:11876594
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Bandman, O., Hillman, J.L., Tang, Y.T., Lal, P., Yue, H., Baughn, M.R.,
 Lu, D.A. and Azimzai, Y.
 TITLE Extracellular matrix and adhesion-associated proteins
 JOURNAL Patent: WO 0068380-A 34 16-NOV-2000;
 Incyte Genomics, Inc. (US)
 FEATURES
 Location/Qualifiers
 1. .1215
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 2040679CB1"

ORIGIN

Query Match 100.0%; Score 1215; DB 6; Length 1215;
 Best Local Similarity 100.0%; Pred. No. 1.8e-261;
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGAACTAGCATGTATGTTATCTCCAGTGGAAATTTAAATCTCAACATTTTATTA 60
 DB 1 GAAGAACTAGCATGTATGTTATCTCCAGTGGAAATTTAAATCTCAACATTTTATTA 60
 QY 61 TTGAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGCTCATAGAG 120

Db 61 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCATAGAG 120
QY 121 GCTGTAATGTTGATACACAGCTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAAA 189
Db 121 GCTGTAATGTTGATACACAGCTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAAA 189
QY 181 ACTTTAAACCTAAATGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGAAATTTTC 243
Db 181 ACTTTAAACCTAAATGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGAAATTTTC 243
QY 241 CATATTCCTTGGAACTCTTACAGCTTTCTTACTGTGGCTTTGTCGAGTTGATATGCGTA 300
Db 241 CATATTCCTTGGAACTCTTACAGCTTTCTTACTGTGGCTTTGTCGAGTTGATATGCGTA 300
QY 301 TCGTTTGGCTTAAAGAGCTTTAGAAATAGACTTTGAGTCAACCCATATATAAAGCTTC 360
Db 301 TCGTTTGGCTTAAAGAGCTTTAGAAATAGACTTTGAGTCAACCCATATATAAAGCTTC 360
QY 361 CAGCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAACTCTGG 420
Db 361 CAGCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACTGAAATGACAACTCTGG 420
QY 421 AGTCATTTAGTGTAGCTTTGTGTCTTACTTACACTCCAGAGTCACTTCGAGTTTGGACC 480
Db 421 AGTCATTTAGTGTAGCTTTGTGTCTTACTTACACTCCAGAGTCACTTCGAGTTTGGACC 480
QY 481 TCAGCAAGAACAAATCAAGGACCTCCCTGTGAGTTTTCGAGCTTCAGGAACTTAAAGA 540
Db 481 TCAGCAAGAACAAATCAAGGACCTCCCTGTGAGTTTTCGAGCTTCAGGAACTTAAAGA 540
QY 541 ATTTAAACCTCAGCAATAATGAATGACTCAATTTCTTGGCAAGATGAGCACTAATAA 600
Db 541 ATTTAAACCTCAGCAATAATGAATGACTCAATTTCTTGGCAAGATGAGCACTAATAA 600
QY 601 ACTTCGCTTTTGTGAGCAGCTCGAATTAAGCTTCCATTTTTCCTAGTGAATTTAGAA 660
Db 601 ACTTCGCTTTTGTGAGCAGCTCGAATTAAGCTTCCATTTTTCCTAGTGAATTTAGAA 660
QY 661 ATTTATCCTTCAATACTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAATCTTTT 720
Db 661 ATTTATCCTTCAATACTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAATCTTTT 720
QY 721 CAGTAATAAGCTCGAAGCACCATTAACTTTATTTGGAATCTTTGCAAGAACCATATTAC 780
Db 721 CAGTAATAAGCTCGAAGCACCATTAACTTTATTTGGAATCTTTGCAAGAACCATATTAC 780
QY 781 ATAAATAGGAATTCATATGCTCTATATCATTTCCATTCCTTCCATTCCTTCCATTCCTT 840
Db 781 ATAAATAGGAATTCATATGCTCTATATCATTTCCATTCCTTCCATTCCTTCCATTCCTT 840
QY 841 CCGCAAAAATTTGTTTGTGGAAGATCTGTCTGCAACTCTTTTCAATTCAGGAATCTACTA 900
Db 841 CCGCAAAAATTTGTTTGTGGAAGATCTGTCTGCAACTCTTTTCAATTCAGGAATCTACTA 900
QY 901 CCATGAATCTGCAATCTGTGCCCCACATGTTGCTTCTAGTAGAATTTGGTGGTACTG 960
Db 901 CCATGAATCTGCAATCTGTGCCCCACATGTTGCTTCTAGTAGAATTTGGTGGTACTG 960
QY 961 AAGCACCTTATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
Db 961 AAGCACCTTATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
QY 1021 TAAAGTAATGGGTGAGACCAAGAAATTTCAATAACAGACTCAGTTTGGGGTGCATG 1080
Db 1021 TAAAGTAATGGGTGAGACCAAGAAATTTCAATAACAGACTCAGTTTGGGGTGCATG 1080
QY 1081 TATGATTTTCAGGCTCAAAATTTGGAGTAAGGAAGATTTCTGTATCTTCTGCTGGAGAGA 1140
Db 1081 TATGATTTTCAGGCTCAAAATTTGGAGTAAGGAAGATTTCTGTATCTTCTGCTGGAGAGA 1140
QY 1141 GGAATGTATAGTACTCATTTAGTACCTCCAAAGCTTTTATTAACCAATTTTACT 1200

Db 1141 GGAATGTATAGTACTCATTTAGTACCTCCAAAGCTTTTATTAACCAATTTTACT 1200
QY 1201 TTTAAAAAATAAA 1215
Db 1201 TTTAAAAAATAAA 1215
RESULT 3
AR339608 1471 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 1099 from patent US 6569662.
ACCESSION AR339608
VERSION AR339608.1 GI:33726465
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1471)
AUTHORS Tang, Y.-T., Zhou, P. and Drmanac, R. T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 1099 27-MAY-2003;
FEATURES Location/Qualifiers
source 1..1471
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 94.9%; Score 1153.2; DB 6; Length 1471;
Best Local Similarity 99.7%; Pred. No. 1-2e-247;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 117
Db 314 TAAAGTAAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 373
QY 118 GAGGCTGTAATGTTGATACACAGCTTTCACGCTCCACAGTGAGAGCTTCAGAAATTC 177
Db 374 GAGGCTGTAATGTTGATACACAGCTTTCACGCTCCACAGTGAGAGCTTCAGAAATTC 433
QY 178 AAAACCTTTAAACTAAATGGTTATCAGATCCAAAAAGACTATCTCTAAGTAAGAAAT 237
Db 434 AAAACCTTTAAACTAAATGGTTATCAGATCCAAAAAGACTATCTCTAAGTAAGAAAT 493
QY 238 TTCCATATTCCTTGGAGACTTCAGACTTCCTTCTGTGGCTTGTCCGAGTTGATATGC 297
Db 494 TTCCATATTCCTTGGAGACTTCAGACTTCCTTCTGTGGCTTGTCCGAGTTGATATGC 553
QY 298 GTATGCTTTGCTTAAAGAGCTTAGGAATTAGACTTTAGTTCACACCATATAAAAAAGC 357
Db 554 GTATGCTTTGCTTAAAGAGCTTAGGAATTAGACTTTAGTTCACACCATATAAAAAAGC 613
QY 358 TTCCAGCTTCAATTTGAGAGCTTCATACACCTTCAAGAACTTAACCTGAAATGCAATCACT 417
Db 614 TTCCAGCTTCAATTTGAGAGCTTCATACACCTTCAAGAACTTAACCTGAAATGCAATCACT 673
QY 418 TGGAGTCAATTTAGTGTGAGCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 477
Db 674 TGGAGTCAATTTAGTGTGAGCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 733
QY 478 ACCTCAGCAAGAACAAAAATCAAGGCACTCCCTGTGCTGAGTTTGGCCAGCTCCAGGAACCTA 537
Db 734 ACCTCAGCAAGAACAAAAATCAAGGCACTCCCTGTGCTGAGTTTGGCCAGCTCCAGGAACCTA 793
QY 538 AGAATTTAAACTTGACGATATGAATTTGAATTTCAATTTCTTCTGCAAGATAGGCAACTAA 597
Db 794 AGAATTTAAACTTGACGATATGAATTTGAATTTCAATTTCTTCTGCAAGATAGGCAACTAA 853
QY 598 TAAACCTTTCGGCTTTTGTGTCAGCAGCTCCGAATTAAGCTTCCATTTTCTTCTTCTTCT 657
Db 854 TAAACCTTTCGGCTTTTGTGTCAGCAGCTCCGAATTAAGCTTCCATTTTCTTCTTCTTCT 913
QY 658 GAAATTTATCCCTTGAATTAATTTGATCTTTTGGAAATACTTTTGAACACCAAAAGTCC 717


```

Db      914  GAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATACCTTTTGAACACCAAAAAGTCC 973
QY      718  TTCAGTAATAAGCTGCAAGACACATTAATTTATTTGGAAATCTTCTGCGAAGAACCATAT 777
Db      974  TTCAGTAATAAGCTGCAAGACACATTAATTTATTTGGAAATCTTCTGCGAAGAACCATAT 1033
QY      778  TACATAATAGATTCATATAGGCTCTCATATCAATTCATTCATCCATCTGCGAAGATTTGG 837
Db      1034  TACATAATAGATTCATATAGGCTCTCATATCAATTCATTCATCCATCTGCGAAGATTTGG 1093
QY      838  ATACCGCAAAAATTTGGTTTGGGAAGATCTCTCTGAATCTTTCAATTCAGGAAGACTA 897
Db      1094  ATACCGCAAAAATTTGGTTTGGGAAGATCTCTCTGAATCTTTCAATTCAGGAAGACTA 1153
QY      898  CTACCATGATCTGCATCTGTTGCCACACATGCTGGTCTTAGTAGATAATTTGGTGGGTA 957
Db      1154  CTACCATGATCTGCATCTGTTGCCACACATGCTGGTCTTAGTAGATAATTTGGTGGGTA 1213
QY      958  CTGAAGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTGTATGTAATTTCTCTGATA 1017
Db      1214  CTGAAGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTGTATGTAATTTCTCTGATA 1273
QY      1018  TGTAAAGTAATGGGTGAGACCGAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGC 1077
Db      1274  TGTAAAGTAATGGGTGAGACCGAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGC 1333
QY      1078  ATCTATGATTTTCAGGCTCAATTTGGAGTAAAGGAAGATTTCTGTATCTGCTGAGA 1137
Db      1334  ATCTATGATTTTCAGGCTCAATTTGGAGTAAAGGAAGATTTCTGTATCTGCTGAGA 1393
QY      1138  GGAGGAATGTATAGTTACTCTATTTAGATGACTCCCAAACTTTTATTAACCAATTTT 1197
Db      1394  GGAGGAATGTATAGTTACTCTATTTAGATGACTCCCAAACTTTTATTAACCAATTTT 1453
QY      1198  AGTTTAAAAAAAATAA 1215
Db      1454  AGTTTAAAAAAAATAA 1471

RESULT 4
LOCUS   BC064133
DEFINITION Homo sapiens cDNA clone MGC:74569 IMAGE:6501107, complete cds.
ACCESSION BC064133
VERSION   BC064133.1 GI:39794694
KEYWORDS MGC.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Strausberg, R., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altshul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, J., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
Mckernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1613)

Strausberg, R.
Direct Submission
Submitted (11-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Varon Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC P-rate: 141 Row: 1 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22748718.

FEATURES

source

1. 1613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:74569 IMAGE:6501107"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
70..264
/codon_start=1
/product="Unknown (protein for MGC:74569)"
/protein_id="AAH64133.1"
/db_xref="GI:39794694"
/translation="MKLHCEVEVISRHLPALGLNRNKGVRVAVLSLCQQTSSQPPVR
AFLLISTLXDKRGTRYHSG"

CDS

ORIGIN

Query Match 94.9%; Score 1153.2; DB 9; Length 1613;
Best Local Similarity 99.7%; Pred. No. 1.2e-247;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTATTCAGGCAATTCAGCAGTTTAAAGGTTTCTCTTCAGCTATGAGACTGGCTCAT 117
Db 452 TAAGTAAGGCCATTCAGCAGTTTAAAGGTTTCTCTTCAGCTATGAGACTGGCTCAT 511
QY 118 GAGGCTGTAATGTGTATACACCGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 177
Db 512 GAGGCTGTAATGTGTATACACCGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 571
QY 178 AAAACTTTAAACTAAATGGTTATACATCCAAAGAAAGACTATCTCTTAAGTAAGAAATT 237
Db 572 AAAACTTTAAACTAAATGGTTATACATCCAAAGAAAGACTATCTCTTAAGTAAGAAATT 631
QY 238 TTCCATATTCCTTGGAACTATCTTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
Db 632 TTCCATATTCCTTGGAACTATCTTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 691
QY 298 GTATGCTTGTCTTAAAGGCTTAGGAAATAGACTTGTAGTCAACACCATATAAAAGG 357
Db 692 GTATGCTTGTCTTAAAGGCTTAGGAAATAGACTTGTAGTCAACACCATATAAAAGG 751

TITLE
JOURNAL
PUBMED

QY 358 TTCCAGTCAATTTGAGAGCTTCATACACCTTCAAGAACTTAACTGAATCAATCACT 417
 Db 752 TTCCAGTCAATTTGAGAGCTTCATACACCTTCAAGAACTTAACTGAATCAATCACT 811
 QY 418 TGGAGTCATTTAGTGTAGCTTGTGTATCTTCACTCCAGAACTCACTTGGAGTTGG 477
 Db 812 TGGAGTCATTTAGTGTAGCTTGTGTATCTTCACTCCAGAACTCACTTGGAGTTGG 871
 QY 478 ACCTCAGCAGACAAATCAAGGCACTCCCTGTGAGTTTCCGAGTCCAGAACTTA 537
 Db 872 ACCTCAGCAGACAAATCAAGGCACTCCCTGTGAGTTTCCGAGTCCAGAACTTA 931
 QY 538 AGAATTTAAACTTTGAGCAATGAATTCATCAATTTCTTCAAGATAGGCAACTAA 597
 Db 932 AGAATTTAAACTTTGAGCAATGAATTCATCAATTTCTTCAAGATAGGCAACTAA 991
 QY 598 TAAACCTTGCCTTTTGTGAGAGCTCGAAATAGCTTCAATTTTGCCTAGTGAATTTA 657
 Db 992 TAAACCTTGCCTTTTGTGAGAGCTCGAAATAGCTTCAATTTTGCCTAGTGAATTTA 1051
 QY 658 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGAACACCAAAAGTCC 717
 Db 1052 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGAACACCAAAAGTCC 1111
 QY 718 TTCCAGTCAATTTGAGCAATGAATTCATCAATTTCTTCAAGATAGGCAACTAA 777
 Db 1112 TTCCAGTCAATTTGAGCAATGAATTCATCAATTTCTTCAAGATAGGCAACTAA 1171
 QY 778 TACATAATAGGATTCATATGGCTCTCATATCATCTCCATTCCTGCAAGATTTGG 837
 Db 1172 TACATAATAGGATTCATATGGCTCTCATATCATCTCCATTCCTGCAAGATTTGG 1231
 QY 838 ATACCGCAAAATTTGTGTTTGGAGAGTCTGTCTGAATCTTTCAATTCAGGAACTA 897
 Db 1232 ATACCGCAAAATTTGTGTTTGGAGAGTCTGTCTGAATCTTTCAATTCAGGAACTA 1291
 QY 898 CTACATGATCTGATCTGTGTTGAGCAATGAATTCATCAATTCAGGAACTA 957
 Db 1292 CTACATGATCTGATCTGTGTTGAGCAATGAATTCATCAATTCAGGAACTA 1351
 QY 958 CTGAGCAGCTATATCTTCTTCTTCTGTTCTAGGCTGTATGTTTCTCTGATA 1017
 Db 1352 CTGAGCAGCTATATCTTCTTCTTCTGTTCTAGGCTGTATGTTTCTCTGATA 1411
 QY 1018 TGTAAAGTAATGGGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTGGGTTG 1077
 Db 1412 TGTAAAGTAATGGGTGAGCAGCAAGAAAGAAATTTCAATAACAGATCAGTTGGGTTG 1471
 QY 1078 ATGTATGATTTTGCAGCGTCAATTTGGAGTAAGGAGATTTCTGTATCTTCTGAGA 1137
 Db 1472 ATGTATGATTTTGCAGCGTCAATTTGGAGTAAGGAGATTTCTGTATCTTCTGAGA 1531
 QY 1138 GGAGGAATGTGTATAGTACTCAATTTGATGATCTCCAAATCTTTTATTAACCAATTTT 1197
 Db 1532 GGAGGAATGTGTATAGTACTCAATTTGATGATCTCCAAATCTTTTATTAACCAATTTT 1591
 QY 1198 AGTTTAAAAAATAAAA 1215
 Db 1592 AGTTTAAAAAATAAAA 1609

RESULT 5
 AY052405
 LOCUS AY052405 Homo sapiens 1445 bp mRNA linear PRI 10-OCT-2001
 DEFINITION Homo sapiens 4-1BB-mediated signaling molecule mRNA, complete cds.
 ACCESSION AY052405
 VERSION AY052405.1 GI:16024937
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1445)

AUTHORS Kwon, B.S. and Jang, I.K.
 TITLE Characterization of LRR-1, a new leucine-rich repeat-containing protein involved in 4-1BB-mediated signal transduction
 JOURNAL Mol. Cells (2001) In press
 REFERENCE 2 (bases 1 to 1445)
 AUTHORS Kwon, B.S. and Jang, I.K.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-2001) Ophthalmology, Louisiana State University, 2020 Gravier Street, New Orleans, LA 70112, USA

FEATURES
 source
 1..1445
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 11..1255
 /note="4-1BBLrr; LRR-1 signaling protein"
 /codon_start=1
 /product="4-1BB-mediated signaling molecule"
 /protein_id="AAU11430.1"
 /db_xref="GI:16024938"
 /translation="MKLHCEVEVISRHLPALGRNKRKVRVSLCQQTSSQPPVR
 AFLISTLKDNRGTRVLERENIQFFTFDEKATVRLKEPVDICLSAISSLKQ
 FLSAMEIAHRGCVNDTPVSTLTPVKTSEFENFKTQWITTSKDYPLSKNPSLSHQ
 TSCGIVRDMRMCLKSLRKLDSLNHIKKLPATIGDLIELOELNLDNHLSPFSA
 LCHSTLQKSLRSLDLSKNIKALPVQCOLQELKNLKLDDNELIOPKCKGOLINLR
 LSAARNKLPLPSEFNLSELYDLRGNTFEQKVLPIKLAFLILLESARKTILHN
 RPYGSHIIPFHLCDLDTAKICVGRFCLNSFIQGTITTNLHSHVAETVWLVDLGGT
 EAPITSYFCSLGCYVNSDMLK"

ORIGIN
 Query Match 93.7%; Score 1138; DB 9; Length 1445;
 Best Local Similarity 99.5%; Pred. No. 3e-244; 5; Indels 1; Gaps 2;
 Matches 1152; Conservative 0; Mismatches

QY 58 TTATT CAGGCATTTCCAGCAGTTTAAAGGTTTCTTTTCCAGCTATCAGACTGGCTCATA 117
 Db 285 TAAGTAAGGCATTTCCAGCAGTTTAAAGGTTTCTTTTCCAGCTATCAGACTGGCTCATA 344
 QY 118 GAGGCTGTAATGTGTATACACGAGTTTCAAGCTGTACACAGTGAAGACTTGCAGACTTG 177
 Db 345 GAGGCTGTAATGTGTATACACGAGTTTCAAGCTGTACACAGTGAAGACTTGCAGACTTG 404
 QY 178 AAAACCTTTAAACCTAAATGGTTTATCACATCCAAAAAGAACTATCTCTTAAGTAAGAAAT 237
 Db 405 AAAACCTTTAAACCTAAATGGTTTATCACATCCAAAAAGAACTATCTCTTAAGTAAGAAAT 464
 QY 238 TTCCATATTTCTTGAACATCTTACAGACTTCTTACCTGTGGGCTTGTCCGAGTTGATATGC 297
 Db 465 TTCCATATTTCTTGAACATCTTACAGACTTCTTACCTGTGGGCTTGTCCGAGTTGATATGC 524
 QY 298 GTATGCTTTTGTAAAGGCTTAGGAAATTTAGACTTTGAGTGCACACCATATATAAAGGC 357
 Db 525 GTATGCTTTTGTAAAGGCTTAGGAAATTTAGACTTTGAGTGCACACCATATATAAAGGC 584
 QY 358 TTCCAGCTCAATTTGAGACCTCATACACCTTCAAGAACTTAACTGAATGACATCACT 417
 Db 585 TTCCAGCTCAATTTGAGACCTCATACACCTTCAAGAACTTAACTGAATGACATCACT 644
 QY 418 TGGAGTCATTTAGTGTAGCTTGTGTATCTTACACTCCAGAGTCTCTCGAGTTTGG 477
 Db 645 TGGAGTCATTTAGTGTAGCTTGTGTATCTTACACTCCAGAGTCTCTCGAGTTTGG 704
 QY 478 ACTCAGCAAGAAACAAATCAAGGCATCTCCCTGTGCTGAGTTTTCAGAGTCCAGGAACTTA 537
 Db 705 ACTCAGCAAGAAACAAATCAAGGCATCTCCCTGTGCTGAGTTTTCAGAGTCCAGGAACTTA 764
 QY 538 AGAATTTAAACTTGAAGATAATGAATTTGATCAATTTCTTGAAGATAGGCAACTAA 597
 Db 765 AGAATTTAAACTTGAAGATAATGAATTTGATCAATTTCTTGAAGATAGGCAACTAA 824
 QY 598 TAAACCTTGCCTTTTGTGAGAGCTCGAAATAGCTTCAATTTTGCCTAGTGAATTTA 657
 Db 825 TAAACCTTGCCTTTTGTGAGAGCTCGAAATAGCTTCAATTTTGCCTAGTGAATTTA 884

```

658 GAAATTTATCCCTTGAATCTGGATCTTTTGGAAATACCTTTGAACAAACCAAGTCC 717
Dbb|||||
885 GAAATTTATCCCTTGAATCTGGATCTTTTGGAAATACCTTTGAACAAACCAAGTCC 944
Qy|||||
718 TTCCAGTAAATAAGCTGCAAGCACCACATTAATCTTTATTTGGAATCTTTCTGCGAAGACCATAT 777
Dbb|||||
945 TTCCAGTAAATAAGCTGCAAGCACCACATTAATCTTTATTTGGAATCTTTCTGCGAAGACCATAT 1004
Qy|||||
778 TACATAATAGATTCATATAGCTCTCATATCAATTCATCTCCATCTCTGCGAAGTGG 837
Dbb|||||
1005 TACATAATAGATTCATATAGCTCTCATATCAATTCATCTCCATCTCTGCGAAGTGG 1064
Qy|||||
838 ATACCGCAAAAATTTGGTTTGGTGAAGATCTCTCTGAACTCTTTTCAATCAAGGACTA 897
Dbb|||||
1065 ATACCGCAAAAATTTGGTTTGGTGAAGATCTCTCTGAACTCTTTTCAATCAAGGACTA 1124
Qy|||||
898 CTACCATGAATCTGCAATCTGTTGCCCCACACCTGCTGTCTTAGTAGATAAATTTGGTGGTA 957
Dbb|||||
1125 CTACCATGAATCTGCAATCTGTTGCCCCACACCTGCTGTCTTAGTAGATAAATTTGGTGGTA 1184
Qy|||||
958 CTGAGCACCATTATCTCTTAATCTCTGTTCTCTAGGCTGTATTTAAATCTCTCTGATA 1017
Dbb|||||
1185 CTGAGCACCATTATCTCTTAATCTCTGTTCTCTAGGCTGTATTTAAATCTCTCTGATA 1244
Qy|||||
1018 TGTAAAGTAATGGGTGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1077
Dbb|||||
1245 TGTAAAGTAATGGGTGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1304
Qy|||||
1078 ATGTATGATTTTGCAGCGTCAATTTGGAGTGAAGAAATTTCTGTATATCTTGTGGAGA 1137
Dbb|||||
1305 ATGTATGATTTTGCAGCGTCAATTTGGAGTGAAGAAATTTCTGTATATCTTGTGGAGA 1364
Qy|||||
1138 GGAGGAATGTATAGTACTCAATTTAGTAGACTCCCAAACTTTTATTTAAACCAATTTT 1197
Dbb|||||
1365 GGAGGAATGTATAGTACTCAATTTAGTAGACTCC-AAACTTTTATTTAAACCAATTTT 1423
Qy|||||
1198 AGTTTAAATAAAAAAAA 1215
Dbb|||||
1424 AGTTTAAATAAAAAAAA 1441

```

```

RESULT 6
LOCUS HSM808175
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J1525 (from clone DKFZp686J1525).
ACCESSION BX648029
VERSION BX648029.1 GI:34367188
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1950)
AUTHORS Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Wewes, H.W., Neill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686J1525) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..1950

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J1525"
/tissue_type="human testis"
/clone_lib="686 (synonym: hlec3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polya_signal 1867..1872
polya_site 1887
ORIGIN
Query Match 81.2%; Score 386.8; DB 9; Length 1950;
Best Local Similarity 88.1%; Pred. No. 2e-210;
Matches 1151; Conservative 0; Mismatches 2; Indels 153; Gaps 1;
Qy 63 CAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTTTCAGCTATGAGACTGGCTCATAGAGGC 122
Dbb|||||
594 CAGGCCAATTTCCAGCAGTTTAAAGGTTTCCTTTTCAGCTATGAGACTGGCTCATAGAGGC 653
Qy|||||
123 TGTAAATGTTGATACACCAAGTTTCAACGCTCACACCAGTGAAGACTTTCAGAAATTTGAAAC 182
Dbb|||||
654 TGTAAATGTTGATACACCAAGTTTCAACGCTCACACCAGTGAAGACTTTCAGAAATTTGAAAC 713
Qy|||||
183 TTTAAACTAAATGGTTTATCATCTCAAAAAGACTATCTCTTAAGTAAGAAATTTTCCA 242
Dbb|||||
714 TTTAAACTAAATGGTTTATCATCTCAAAAAGACTATCTCTTAAGTAAGAAATTTTCCA 773
Qy|||||
243 TATTCCTTGAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATCGGTATG 302
Dbb|||||
774 TATTCCTTGAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATCGGTATG 833
Qy|||||
303 CTTTGCTTAAAGGCTTTAGGAAATTAAGCTGAGTCAACCAATATAAAAGAACTTTCCA 362
Dbb|||||
834 CTTTGCTTAAAGGCTTTAGGAAATTAAGCTGAGTCAACCAATATAAAAGAACTTTCCA 893
Qy|||||
363 GCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACTGGAG 422
Dbb|||||
894 GCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACTGGAG 953
Qy|||||
423 TCATTTAGTGTAGCTTTGTCTCATTTCTACACTCCAGAACTCACTTCGAGAGTTTGGACCTC 482
Dbb|||||
954 TCATTTAGTGTAGCTTTGTCTCATTTCTACACTCCAGAACTCACTTCGAGAGTTTGGACCTC 1013
Qy|||||
483 AGCAAGAACAAATCAAGGACCTCCCTCTGCACTTTTCCAGCTCCAGGAACTTAAGAT 542
Dbb|||||
1014 AGCAAGAACAAATCAAGGACCTCCCTCTGCACTTTTCCAGCTCCAGGAACTTAAGAT 1073
Qy|||||
543 TTAATAACTTGAACGATAATGAATTTGATTTCAATTTCTTTCAGAGATAGACAACTAATAAC 602
Dbb|||||
1074 TTAATAACTTGAACGATAATGAATTTGATTTCAATTTCTTTCAGAGATAGACAACTAATAAC 1133
Qy|||||
603 CTTTCGCTTTTGTGACAGCTCGAAATAAGACTTCCATTTTTCGCTAGTGAATTTAGAAAT 662
Dbb|||||
1134 CTTTCGCTTTTGTGACAGCTCGAAATAAGACTTCCATTTTTCGCTAGTGAATTTAGAAAT 1193
Qy|||||
663 TTATCCCTTGAATACACTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCCCTTCCA 722
Dbb|||||
1194 TTATCCCTTGAATACACTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCCCTTCCA 1253
Qy|||||
723 GTAATAAAGCTGCAAGCACCATTAACTTTATTTGGAATCTTCTGCGAAGACCATATTACAT 782
Dbb|||||
1254 GTAATAAAGCTGCAAGCACCATTAACTTTATTTGGAATCTTCTGCGAAGACCATATTACAT 1313
Qy|||||
783 AAT----- 785
Dbb|||||
1314 AATAGAACCATCTATCCGAAGAGATTAATGAAACATGTGGAGAGACTCCCTCGGAGCA 1373
Qy|||||
786 ----- 785
Dbb|||||
1374 GTTTTAAAAAATGGTTACTTAAGCAACAGATTAAAGTCTTTTATGAAAAATTTAA 1433
Qy|||||
786 -----AGGATTCATATGGCTCTCATATC 809

```

Db 1434 AAATTGCAATAGGCCCATCTAGATCCCGTACCAAGAGTTCATATGGCTCTCATATC 1493
QY 810 ATTCCATTCCATCTCTGCCAAGATTTCGATACCGCAAAATTTTGTGTTGGAGATTC 869
Db 1494 ATTCCATTCCATCTCTGCCAAGATTTCGATACCGCAAAATTTTGTGTTGGAGATTC 1553
QY 870 TGTCTGAATCTTTCAATCAAGGAATCTACCATGAATTCGATTCCTGTTGCCCACT 929
Db 1554 TGTCTGAATCTTTCAATCAAGGAATCTACCATGAATTCGATTCCTGTTGCCCACT 1613
QY 930 GTGGCTTTAGTAGATAATTTGGGTGGTACTGAGACCACTTAATCTCTTATTTCTGTTCT 989
Db 1614 GTGGCTTTAGTAGATAATTTGGGTGGTACTGAGACCACTTAATCTCTTATTTCTGTTCT 1673
QY 990 CTAGGCTGTTATCTTAATCTCTGATCTGTAATGTAAGTAAGTGGTGAGACCAAGAA 1049
Db 1674 CTAGGCTGTTATCTTAATCTCTGATCTGTAATGTAAGTGGTGAGACCAAGAA 1733
QY 1050 ATTCAATAACAGATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAATTTGGAGTAA 1109
Db 1734 ATTCAATAACAGATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAATTTGGAGTAA 1793
QY 1110 GGGAGATTTCTGTATCTGTCATCTGTCGAGAGGAGAACTGTGTATAGTACTCAATTAGATGA 1169
Db 1794 GGGAGATTTCTGTATCTGTCATCTGTCGAGAGGAGAACTGTGTATAGTACTCAATTAGATGA 1853
QY 1170 CTCCTAAATCTTTATTAACCAATTTTGTAGTTTAAAAA 1215
Db 1854 CTCCTAAATCTTTATTAACCAATTTTGTAGTTTAAAAA 1899

RESULT 7
CNS01DX7/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CNS01DX7
Human chromosome 14 DNA sequence BAC R-649E7 of library RPCI-11
from chromosome 14 of Homo sapiens (human), complete sequence.
AL139099
AL139099.3 GI:15424523
HTG: HTGS ACTIVEFIN.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157010)
Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissenbach, J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 157010)
Genoscope.
Direct Submission
Submitted (31-AUG-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr;
On Sep 3, 2001 this sequence version replaced gi:8217934.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Seqref@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-831F12 (AC=AL591767)
Downstream BAC (overlapping the SP6 end) : R-346L24 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.30x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases
0 :
1 - 9 : 8
10 - 19 : 92
20 - 29 : 225
30 - 39 : 870
40 - 49 : 3443
50 - 59 : 7315
60 - 69 : 11031
70 - 79 : 25134
80 - 89 : 50478
90 - 99 : 58414
Percentage of bases with a quality value >= 40 : 99 %.

location/Qualifiers
1. 157010
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone_lib="RPCI-11"
31812..32022
/note="matching EMBL:Z52356
RDB:RH31309
dbSTS:STS15763
Identified using the e-PCR software (G. Schuler)"
/note="matching EMBL:N22189
RDB:RH69276
dbSTS:STS49154
Identified using the e-PCR software (G. Schuler)"
41035..41109
/note="matching EMBL:G31448
RDB:RH41760
dbSTS:STS35761
Identified using the e-PCR software (G. Schuler)"
34622..43745
/note="matching EMBL:M15661
RDB:RH71370
dbSTS:STS51135
Identified using the e-PCR software (G. Schuler)"
48088..48262
/note="matching EMBL:AA424400
RDB:RH103208
dbSTS:STS70744
Identified using the e-PCR software (G. Schuler)"
64270..64448
/note="matching EMBL:W84670
RDB:RH82026
dbSTS:STS58378
Identified using the e-PCR software (G. Schuler)"

ORIGIN
Query Match 64.9%; Score 788.6; DB 9; Length 157010;
Best Local Similarity 98.3%; Pred. No. 2.5e-166;
Matches 797; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GAAGACTAGCATGTATGTTATCTCCAGTGAATTTATAATTTTACACTTTTATTA 60
Db 55298 GAAGAACTAGCATGTATGTTATCTCCAGTGAATTTATAATTTTACACTTTTATTA 55239
QY 61 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGCTGGCTCATAG 120
Db 55238 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGCTGGCTCATAG 55179
QY 121 GCTGTAATGTTGATACACCACTTTCAACGCTCACACCACTGAGACTTCAGAAATTTGAAA 180
Db 55178 GCTGTAATGTTGATACACCACTTTCAACGCTCACACCACTGAGACTTCAGAAATTTGAAA 55119
QY 181 ACTTTAAACTAAATGTTATCATCATCCAAAAAGACTATCTCTTAAGTAAGATTTTC 240
Db 55118 ACTTTAAACTAAATGTTATCATCATCCAAAAAGACTATCTCTTAAGTAAGATTTTC 55059

```

QY 241 CATATTCCTTGGAAATCTTTGAGACTCTTACTGTGGGCTTGTGGAGTTGATATGCGTA 300
D 55058 CATATTCCTTGGAAATCTTTGAGACTCTTACTGTGGGCTTGTGGAGTTGATATGCGTA 54999
QY 301 TGCCTTCTTAAAGAGCTTTAGGAAATAGACTTGGAGTCAACCAATATAAAAAAGCTTC 360
D 54998 TGCCTTCTTAAAGAGCTTTAGGAAATAGACTTGGAGTCAACCAATATAAAAAAGCTTC 54999
QY 361 CAGTACAAATGGAGAGCTCATACACCTTCAAGAACTTAACCTGAATGACCAATCATTGG 420
D 54938 CAGTACAAATGGAGAGCTCATACACCTTCAAGAACTTAACCTGAATGACCAATCATTGG 54879
QY 421 AGTCATTTAGTGTAGCTTGTGTCTCACTACCTCCAGAGTCACTTCGGAGTTTGGACC 480
D 54878 AGTCATTTAGTGTAGCTTGTGTCTCACTACCTCCAGAGTCACTTCGGAGTTTGGACC 54819
QY 481 TCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCAGCTCCAGGAATTAAGA 540
D 54818 TCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCAGCTCCAGGAATTAAGA 54759
QY 541 ATTATTAAGCTTGGAGTAAATGAATGATTCATTCCTCCAGATGAGGACCACTAATAA 600
D 54758 ATTATTAAGCTTGGAGTAAATGAATGATTCATTCCTCCAGATGAGGACCACTAATAA 54699
QY 601 ACCTTCGCTTTTGTGTCAGAGCTCGAATAAGCTTCCATTTTTCCTAGTGAATTTAGAA 660
D 54698 ACCTTCGCTTTTGTGTCAGAGCTCGAATAAGCTTCCATTTTTCCTAGTGAATTTAGAA 54639
QY 661 ATTATTCCTTGAATCTTGGATCTTTTTCGAAATACCTTTTGAACCAACCAAGTCTCTTC 720
D 54638 ATTATTCCTTGAATCTTGGATCTTTTTCGAAATACCTTTTGAACCAACCAAGTCTCTTC 54579
QY 721 CAGTAATAAGCTGCAAGCACTTATTCCTTATTCGATCTTCTGCAAGCACTATATAC 780
D 54578 CAGTAATAAGCTGCAAGCACTTATTCCTTATTCGATCTTCTGCAAGCACTATATAC 54519
QY 781 ATAATAGGATTCATATGGCTCTCATATCAT 811
D 54518 ATAATAGGATTCATATGGCTCTCATATCAT 54488

RESULT 8
AC126994
LOCUS
DEFINITION Rattus norvegicus clone CH230-295C20, *** SEQUENCING IN PROGRESS ***
ACCESSION AC126994
VERSION AC126994.3 GI:25139354
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 193800)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

```

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschew, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Morris, S., Naokeleneh, O., Okwono, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Snajs, D., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villabana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, J., Zhou, S., Yen, J., Yoon, L., Yoon, V., Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 193800)
Morley, K. C.

Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 193800)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:22855985. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G20F
Center clone name: CH230-235C20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 178664 bases at least Q40
Consensus quality: 181447 bases at least Q30
Consensus quality: 182574 bases at least Q20
Estimated insert size: 188087; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.bgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces
 is believed to be correct as given, however the sizes
 of the gaps between them are based on estimates that have
 been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 193800: contig of 193800 bp in length.

FEATURES

Source

Location/Qualifiers
 1..193800
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-295C20"
 1..2415
 /note="wgs_contig"
 6416..9345
 /note="wgs_contig"
 188582..190216
 /note="wgs_contig"
 191015..193800
 /note="wgs_contig"

ORIGIN

Query Match 46.8%; Score 568.6; DB 2; Length 193800;
 Best Local Similarity 83.3%; Pred. No. 4.1e-117;
 Matches 659; Conservative 0; Mismatches 129; Indels 3; Gaps 1;
 1 GAAGAACTAGCATGTATGTATATCTCCAGTGGAAATTAATTTACAA---CTTTTAT 57
 46535 GAAGAACTAGCAAGCGTGAAGTAATTAAGGTGGACCTTATAATTTGTATTTT 46594
 58 TTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGGCTCATA 117
 46595 TTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGGCTCATA 46654
 118 GAGGCTGTATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGACTTCAGAAATTG 177
 46655 GAGGCTGTATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGACTTCAGAAATTG 46714
 178 AAAATTTTAAACTAAATGTTATCATATCCAAAAGAGACTATCTCTAAGTAGAAT 237
 46715 AAAATTTTAAACTAAATGTTATCATATCCAAAAGAGACTATCTCTAAGTAGAAT 46774
 238 TTCCATATTCCTTGGAAATCTTCAGACTTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
 46775 TTCCATATTCCTTGGAAATCTTCAGACTTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 46834
 298 GTATGCTTTGCTTAAAGGCTTAGGAATATGAGCTTGAAGTACATCAACCATATTAAGAAGC 357
 46835 GTATGCTTTGCTTAAAGGCTTAGGAATATGAGCTTGAAGTACATCAACCATATTAAGAAGC 46894
 358 TTCCAGCTTACAAATTTGGAGACCTTCATACACCTTCAAGAACTTAACCTGAATGACATCACT 417
 46895 TTCCAGCTTACAAATTTGGAGACCTTCATACACCTTCAAGAACTTAACCTGAATGACATCACT 46954
 418 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTTACATCCAGAACTCACTTCGGAGTTGG 477
 46955 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTTACATCCAGAACTCACTTCGGAGTTGG 47314
 478 ACCTTCAGCAAGCAAAATCAAGGACTCCCTGTGAGTTTTCAGCTTCAGGAACTTA 537
 47015 ATCTTCAGCAAGCAAAATCAAGGACTCCCTGTGAGTTTTCAGCTTCAGGAACTTA 47074
 538 AGAATTTTAAACTTGAACATATGATTTGATTTCAATTTCTTGGCAAGATAGCAACTAA 597
 47075 CTAATTTTAAACTTGAACATATGATTTGATTTCAATTTCTTGGCAAGATAGCAACTAA 47134

QY 598 TAAACCTTCGCTTTTGTGACGAGCTCGAAATAGCTTCCATTTTTCGCTAGTAGTAATTA 657
 Db |||||
 47135 CAAACCTTCGCTTTTGTGACGAGCTCGAAATAGCTTCCATTTTTCGCTAGTAGTAATTA 47194
 QY 653 GAAATTTATCCCTTGAATATCTTGATCTTTTGGAAATACCTTTTGAACACCAAGAGTCC 717
 Db |||||
 47195 AATGTTATCCCTTGAATATCTTGATCTTTTGGAAATACCTTTTGAACACCAAGAGTCC 47254
 QY 718 TTCCAGTATTAAGCTGCAAGCACCATTAACTTTTATGGAATCTTCTGACGAACTATAT 777
 Db |||||
 47255 TTCCAGTATTAAGCTGCAAGCACCATTAACTTTTATGGAATCTTCTGACGAACTATAT 47314
 QY 778 TACATTAATAGG 788
 Db |||||
 47315 TATCCTATAGG 47325

RESULT 9

AC099934

LOCUS

DEFINITION

AC099934

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkatachalam, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 17, 2003 this sequence version replaced gi:38502455.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WfBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13587
Center clone name: L1_K_9

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 163083: contig of 163083 bp in length
* 163084 163183: gap of 100 bp
* 163184 232458: contig of 69275 bp in length.

FEATURES

source
Location/Qualifiers
1. .232458
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/map="12"
/clone="RP23-17K9"
/clone_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 46.6%; Score 566; DB 2; Length 232458;
Best Local Similarity 83.1%; Pred. No. 1.5e-116;
Matches 658; Conservative 0; Mismatches 130; Indels 4; Gaps 1;
QY 1 GAAGAACTAGCATGTATGTTATCTCCAGTGGAAATTAATTAATCTACAAAC---TTTAA 56
DB 82575 GGAGAACTAGCAAGCATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 82634
QY 57 TTATTTCCAGGCAATTTCCAGCACTTTAAAGGTTCTTTTCCAGTATGAGACGTGGCTCAT 116
DB 82635 TTATTTCCAGGCAATTTCCAGCACTTTAAAGACTTTACTTTCCAGTATGAGACGTGGCTCAT 82694
QY 117 AGAGGCTGTAATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGCTTCCAGAAATTT 176
DB 82695 AGAGGCTGTAATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGCTTCCAGAAATTT 82754
QY 177 GAAACCTTTAAACCTAAATGTTATACATCCAAAGAAAGACTATCTCTTAAGTAAGAAAT 236
DB 82755 GAAATATATAAACAACCAATGTTATACATCCAAAGAAATTAATTTCTTAACCAAGAAC 82814
QY 237 TTTCATATCTTGGAAATCTTCAGACTTCTTAATCTGTGGGCTTGTTCGAGTTGATATG 296
DB 82815 TTTCATATATTTCTGGAACACCTTCAGGCTTCTTAATCTGTAGCTTGTCCGAGTTGATATG 82874
QY 297 CGTATGCTTTGCTTAAAGACCTTACGAATTAAGCTTGTAGTCACAAACCATATAAAGAG 356

Db 82875 CGTATGCTGCTTTAAAGAAATCTTACGAGTTAGACCTGAGTCTCACTATATAAAAAA 82934
QY 357 CTTCCAGCTACAATTTGGAGACCTCATACACCTTCAAGAACTTAAACCTGAATGACAACTAC 416
DB 82935 CTTCCAGCTACAATTTGGAGACCTCACCCACCTTCAAGAGCTTAACTTAAAGCAATCAG 82994
QY 417 TTGGAGCTATTAGTGTAGCTTGTGTCATCTCTACACTCCAGAGTCACTTGGGAGTTTG 476
DB 82995 CTGGAGACCTTTAGTGTGCTTGTGACTTCCACACTCCAGAGTCACTTCCACAGTTTG 83054
QY 477 GACCTCAGCAGAAACAAATCAAGCCACTCCCTGTGCAAGTTTGGCAGCTCCAGCAACTT 536
DB 83055 GATCTCAGCAGAAACAAATCAAGCGCTCCCTGTGCAAGTTTGTCAATTCGGGAACTT 83114
QY 537 AAGAAATTAACCTGACGATTAATGAATTTCAATTTCCCTGCAAGATAGGCAACTA 596
DB 83115 ACTAACTTAAACCTTAATGAATTAATGAATTTCACTTCCCTTCCAGATAGGACAGCTA 83174
QY 597 ATAAACCTTCGCTTTTGTGACGAGCTCGAATTAAGCTTCCACTTTTGCCTTAGTAATTT 656
DB 83175 ACAACCTTCGCTTTTCTATCAGCAGCTCGAATTAAGCTTCCACTTCCCTCTGAATTT 83234
QY 657 AGAAATTTATCCCTTGAATTAATTTGGATCTTTTGGAAATTAATTTTGAACCAACCAAGTC 716
DB 83235 AAAATGTTATCCCTGGAGTACTTGGATCTTTTGGAAATTAATTTTGAACCAACCAAGTC 83294
QY 717 CTTCCAGTATAAAGCTGCAAGCACTTAATTAATTTTGAATCTTCTGCAAGCAACTA 776
DB 83295 ATTCCAATTAAGCTTCAAGTACCATACTTACTGGAATCATGTGAACAAGCTGTG 83354
QY 777 TTACATAATAGG 788
DB 83355 CTATCTATAGG 83366

RESULT 10
AX255882
LOCUS AX255882 682 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 33 from Patent WO0170976.
ACCESSION AX255882
VERSION AX255882.1 GI:16074922
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu, J., Pyle, R.A. and Stoik, J.A.
TITLE Compositions and methods for the therapy and diagnosis of ovarian and endometrial cancer
JOURNAL Patent: WO 0170976-A 33 27-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. .682
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 42.9%; Score 521.6; DB 6; Length 682;
Best Local Similarity 97.5%; Pred. No. 2.8e-106;
Matches 55; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY 277 GGCTTCCGAGTTGATATGCGTATGCTTGGCTTAAGAGCCTTAGAAATTAAGCTTGA 336
DB 1 GGCTTCCGAGTTGATATGCGTATGCTTGGCTTAAGAGCCTTAGAAATTAAGCTTGA 60
QY 337 GTCACAAACCATATAAAGAGCTTCCAGTACAAATTCGAGACCTCATACACCTTCAAGAAC 396
DB 61 GTCACAAACCATATAAAGAGCTTCCAGTACAAATTCGAGACCTCATACACCTTCAAGAAC 120
QY 397 TTAACCTGAATGACATGCTGAGTCACTTAGTGTAGCTTGTGTCTATCTACACTCC 456

Db 121 TTAACCTGAATGACAAATCACTTGGAGTCATTTAGTGTAGCCTTGTGTCACTTCACTCC 180
 QY 457 AGAAGTCACCTTCGAGTTTGGACCTCAGCAGAGACAAATCAAGGCACCTCCCTGTGAGT 516
 Db 181 AGAAGTCACCTTCGAGTTTGGACCTCAGCAGAGACAAATCAAGGCACCTCCCTGTGAGT 240
 QY 517 TTTGCCAGCTCCAGGAACCTTAAGAACTTTAAAGAACTTTGACGATTAATGAATTCATTTTC 576
 Db 241 TTTGCCAGCTCCAGGAACCTTAAGAACTTTAAAGAACTTTGACGATTAATGAATTCATTTTC 302
 QY 577 CTTCGAGATAGGCAACTTAATAAACCCTTCCTTTTGTCTAGCAGCTCGAAATTAAGCTTC 636
 Db 301 CTTCGAGATAGGCAACTTAATAAACCCTTCCTTTTGTCTAGCAGCTCGAAATTAAGCTTC 360
 QY 637 CATTTTTCCTAGTGAATTTAGAAATTTTCCCTTGTGATTAATCTTGGATCTTTTGGAAATA 696
 Db 361 CATTTTTCCTAGTGAATTTAGAAATTTTCCCTTGTGATTAATCTTGGATCTTTTGGAAATA 420
 QY 697 CTTTGAACCAACCAAAAGTCCTTCCAGTAATAAGCTTGAAGCTGCAAGCACCAATTAATTTTGG 756
 Db 421 CTTTGAACCAACCAAAAGTCCTTCCAGTAATAAGCTTGAAGCTGCAAGCACCAATTAATTTTGG 480
 QY 757 AATCTTCTGACGAACCAATTAATTAATTAAT--AGATTCATATGCTCT--CATATCAT 812
 Db 481 AATCTTCTGACGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 QY 813 CCATTCATCTCTGCCAAGATTTGG 837
 Db 541 CCATTCATCTCTGCCAAGATTTGG 565

RESULT 11
 AX336594/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AX336594
 Sequence 7103 from Patent WO0194629.
 AX336594
 AX336594.1 GI:18127313
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES

Young, P.E., Augustus, X., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 7103 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 Location/Qualifiers
 1..467
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 36.5%; Score 443.8; DB 6; Length 467;
 Best Local Similarity 98.1%; Pred. No. 7.6e-89;
 Matches 460; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 739 CACCAATTAATTTTGGAACTTCTTGGCAGCAACCATATTACATATAGGATTCATATG 798
 Db 467 CACCAATTAATTTT--GGAATCTTGGCAGCAACCATATTACATATAGGATTCATATG 410
 QY 799 GCTCTCATATCATTCATTCATTCCTCTGCAAGATTTGGATCCGCAAAATTTGGTTT 858
 Db 409 GCTCTCATATCATTCATTCATTCCTCTGCAAGATTTGGATCCGCAAAATTTGGTTT 350
 QY 859 GTGGAAGATTTCTGTGAACTTTTCAATTCAGGAACCTACTACCATGATTCGATTCG 918
 Db 349 GTGGAAGATTTCTGTGAACTTTTCAATTCAGGAACCTACTACCATGATTCGATTCG 290
 QY 919 TTGCCACACTGTGCTCTTAGTAGATAATTTGGTGGTACTGAAGCACCTATATCTCTT 978

Db 289 TTGCCACACCTGTGCTTAGTAGATAATTTGGTGGTACTGAAGCACCTATATCTCTT 230
 QY 979 ATTCTGTCTCTAGGCTGTATGTAATTCCTCTGATATGTTAAAGTAATGGTGAGAC 1038
 Db 229 ATTCTGTCTCTAGGCTGTATGTAATTCCTCTGATATGTTAAAGTAATGGTGAGAC 170
 QY 1039 CAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCATGTATGATTTGCACGCTCA 1098
 Db 169 CAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCATGTATGATTTGCACGCTCA 110
 QY 1099 AATTGGAGTAAGGAAGATTTCTGTACTACTTGTCTGGAGGAGGAATCTGTATGTTACT 1158
 Db 109 AATTGGAGTAAGGAAGATTTCTGTACTACTTGTCTGGAGGAGGAATCTGTATGTTACT 50
 QY 1159 CATTTAGTAGTCTCAAAACCTTTTATTTAAACCAATTTTACTTTTAAAA 1207
 Db 49 CATTTAGTAGTCTCAAAACCTTTTATTTAAACCAATTTTACTTTTAAAA 1

RESULT 12
 BC030142
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BC030142
 Homo sapiens peptidylprolyl isomerase (cyclophilin) like 5, mRNA
 (CDNA clone MGC:20689 IMAGE:4764756), complete cds.
 BC030142
 BC030142.2 GI:33988785
 MGC.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Xarusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.H.,
 Schetz, T.E., Brownstein, M.J., Udell, C.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Li, X., Gibbs, R.A.,
 Fahey, J., Heiton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smal, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257
 12477932
 2 (bases 1 to 990)
 Strausberg, R.
 Direct Submission
 Submitted (07-MAY-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 20, 2003 this sequence version replaced gi:20985783.
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Louis Staudt
 cDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL);
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 6164 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/6164
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT source 1..584
 FT /organism="Homo sapiens (human)".
 FT Location/Qualifiers
 FT 1..584
 FT /organism="Homo sapiens"
 FT /mol_type="genomic DNA"
 FT /db_xref="taxon:9606"

ORIGIN
 Query Match 34.5%; Score 418.8; DB 6; Length 584;
 Best Local Similarity 99.5%; Pred. No. 2.9e-83;
 Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 782 TAATAGGATTCATATGGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 841
 Db 422 TATGAGGATTCATATGGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 363
 QY 842 CGCAAAATTTGTTGTGGAAGATTTCTGTAACCTCTTTCATTCGAGGAATCTAC 901
 Db 362 CGCAAAATTTGTTGTGGAAGATTTCTGTAACCTCTTTCATTCGAGGAATCTAC 303
 QY 902 CATGAATCTGCATTTCTGTCGCCACATCTGCTTAGTAGATTAATTTGGTGTCTGA 961
 Db 302 CATGAATCTGCATTTCTGTCGCCACATCTGCTTAGTAGATTAATTTGGTGTCTGA 243
 QY 962 AGCACTTATATCTTTATTTCTGTTCTAGGCTGTTATGTTAACTTCTCTGATATGTT 1021
 Db 242 AGCACTTATATCTTTATTTCTGTTCTAGGCTGTTATGTTAACTTCTCTGATATGTT 183
 QY 1022 AAGTAATGGTGGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGGTGCATGT 1081
 Db 182 AAGTAATGGTGGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGGTGCATGT 123
 QY 1082 ATGATTTGAGCGCTCAAAATGGAGTAAGGAGATTTCTGTATCTTGTGGAGAGGAG 1141
 Db 122 ATGATTTGAGCGCTCAAAATGGAGTAAGGAGATTTCTGTATCTTGTGGAGAGGAG 63
 QY 1142 GAATGTGTATGATCTATGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1201
 Db 62 GAATGTGTATGATCTATGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 3
 QY 1202 TT 1203
 Db 2 TT 1

RESULT 15
 AX878622
 LOCUS

DEFINITION Sequence 13527 from Patent EP1074617.
 ACCESSION AX878622 GI:40033358
 VERSION AX878622.1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primers for synthesizing full-length cDNA and their use
 Patent: EP 1074617-A 13527 07-FEB-2001;
 Research Association for Biotechnology (JP)
 FEATURES
 source
 1..1590
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 CDS
 385..882
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE90199.1"
 /db_xref="GI:40033359"
 /translation="MSVTLHTDVGDIKIEFCERTPKTCMESRCVPOAGVQWRDLGS
 LOPPRGFKQVFCLSLPRGTGSGNSIWKKFEDEXSEYLKXNVRGVVSMANGPTNG
 SDFEITYGKQPHLDWKYTVFKVLDGLDELKLPVNEKTRPLNDVHLDITIIHA
 NPFAQ"

ORIGIN
 Query Match 34.5%; Score 418.8; DB 6; Length 1590;
 Best Local Similarity 99.5%; Pred. No. 2.5e-83;
 Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 782 TAAAGGATTCATATGGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 841
 Db 1169 TATGAGGATTCATATGGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 1228
 QY 842 CGCAAAATTTGTTGTGGAAGATTTCTGTAACCTCTTTCATTCGAGGAATCTAC 901
 Db 1229 CGCAAAATTTGTTGTGGAAGATTTCTGTAACCTCTTTCATTCGAGGAATCTAC 1288
 QY 902 CATGAATCTGCATTTCTGTCGCCACATCTGCTTAGTAGATTAATTTGGTGTCTGA 961
 Db 1289 CATGAATCTGCATTTCTGTCGCCACATCTGCTTAGTAGATTAATTTGGTGTCTGA 1348
 QY 962 AGCACTTATATCTTTATTTCTGTTCTAGGCTGTTATGTTAACTTCTCTGATATGTT 1021
 Db 1349 AGCACTTATATCTTTATTTCTGTTCTAGGCTGTTATGTTAACTTCTCTGATATGTT 1408
 QY 1022 AAGTAATGGTGGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGGTGCATGT 1081
 Db 1409 AAGTAATGGTGGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGGTGCATGT 1468
 QY 1082 ATGATTTGAGCGCTCAAAATGGAGTAAGGAGATTTCTGTATCTTGTGGAGAGGAG 1141
 Db 1469 ATGATTTGAGCGCTCAAAATGGAGTAAGGAGATTTCTGTATCTTGTGGAGAGGAG 1528
 QY 1142 GAATGTGTATGATCTATGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1201
 Db 1529 GAATGTGTATGATCTATGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1588
 QY 1202 TT 1203
 Db 1589 TT 1590

Search completed: June 16, 2004, 00:15:48
 Job time : 3279 secs